Title

The regulatory and transcriptional landscape associated with carbon utilization in a filamentous fungus.

Journal

Proceedings of the National Academy of Sciences of the United States of America, 117(11)

ISSN

0027-8424

Authors

Wu, Vincent W
Thieme, Nils
Huberman, Lori B
et al.

Publication Date

2020-03-01

DOI

10.1073/pnas.1915611117

The regulatory and transcriptional landscape associated with carbon utilization in a filamentous fungus

Vincent W. Wu[a,b], Nils Thieme[c,g], Lori B. Huberman[a,b], Axel Dietschmann[c,f], David J. Kowbel[a], Juna Lee[d], Sara Calhoun[d], Vasanth Singan[d], Anna Lipzen[d], Yi Xiong[a,b,h], Remo Monti[d], Matthew J. Blow[d], Ronan C. O'Malley[d], Igor V. Grigoriev[d,5], J. Philipp Benz[d], N. Louise Glass[a,b,e,1]

[a]Department of Plant and Microbial Biology, University of California, Berkeley, [b]Energy Biosciences Institute, University of California, Berkeley, [c]Holzforschung München, TUM School of Life Sciences Weihenstephan, Technical University of Munich, Freising, Germany, [d]Joint Genome Institute, Walnut Creek, CA

[e]Environmental Genomics and Systems Biology, Lawrence Berkeley National Laboratory, Berkeley, CA

[f]current address: Chair of Microbiology, TUM School of Life Sciences Weihenstephan, Technical University of Munich, Freising, Germany

[g]current address: Department of Infection Biology, Institute for Clinical Microbiology, Immunology and Hygiene, Universitätsklinikum Erlangen and Friedrich-Alexander Universität, Erlangen-Nürnberg, Germany

[h]current Address: Amyris, Inc. Emeryville, CA

[1]Corresponding Author: N. Louise Glass

Lglass@berkeley.edu

Keywords: Regulatory networks, nutrient signaling, plant biomass, cellulases, pectinases, hemicellulases, DAP-seq, Neurospora, major facilitator transporters, carbon catabolite repression

Classification: Biological Sciences, Genetics

Abstract

Filamentous fungi, such as *Neurospora crassa*, are very efficient in deconstructing plant biomass , both by the secretisecretion ofng an arsenal of plant cell wall degrading enzymes, and by remodeling metabolism to accommodate production of secreted enzymes and by to enablinge transport and intracellular utilization of plant biomass components. Although a number of enzymes and transcriptional regulators involved in plant biomass utilization have been identified, how filamentous fungi sense and integrate nutritional information encoded in the plant cell wall into a regulatory hierarchy for optimal utilization of complex carbon sources is not understood. Here we performed transcriptional profiling of *N. crassa* on 40 different carbon sources, including plant biomass, to provide data on how fungi sense simple to complex carbohydrates. From these data, we identified new regulatory factors in *N. crassa* and characterized one (PDR-2) associated with pectin utilization one with pectin/hemicellulose utilization (ARA-1). Using *in vitro* DNA-affinity purification sequencing (DAP-seq), we identified direct targets of transcription factors involved in regulating genes encoding plant cell wall degrading enzymePCWDEs encoding genes. In particular, our data clarified the role of the transcription factor VIB-1 in the regulation of genes encoding PCDWEs and nutrient scavenging and revealed a major role of the carbon catabolite repressor CRE-1 in regulating the expression of major facilitator transporter genes. These data contribute to a more complete understanding of crosstalk between transcription factors and their target genes, which are involved in regulating nutrient sensing and plant biomass utilization on a global level.

Significance statement

Microorganisms have evolved signaling networks to identify and prioritize utilization of carbon sources. For fungi that degrade plant biomass, such as *Neurospora crassa*, signaling networks dictate the metabolic response to carbon sources present in plant cell walls, resulting in optimal utilization of nutrient sources. However, within a fungal colony, regulatory hierarchies associated with activation of transcription factors and temporal and spatial production of proteins for plant biomass utilization are unclear. Here, we perform expression profiling of *N. crassa* on simple sugars to complex carbohydrates to identify regulatory factors and direct target of regulatory transcription factors using DNA-affinity purification sequencing

(DAP-seq). These findings will enable more precise tailoring of metabolic networks in filamentous fungi for the production of second-generation biofuels.

Introduction

In nature, fungi must integrate acquisition of nutrients with metabolism, growth and reproduction. Fungal deconstruction of plant biomass by fungi requires the ability to efficiently produce and secrete large quantities of secreted plant cell wall degrading enzymes (PCWDEs). Turnover of plant biomass by fungi is an ecosystem function (1), as well as an attribute that has been harnessed industrially to convert plant biomass to simple sugars and in turn high value compounds (2). The plant cell wall is composed of a complex and integrated set of polysaccharides that can vary across tissue type and plant species. Cellulose is the most recalcitrant and most abundant cell wall polysaccharide and is composed of β-1,4-linked D-glucose residues arranged in linear chains. Hemicelluloses represent about 20-35% of primary plant cell wall biomass, and include polysaccharides with β-1,4-linked backbones, such as xylan, xyloglucan and mannan. Pectin is a heterogeneous structure with an abundance of D-galacturonic acid, L-rhamnose and L-arabinose. The two most common forms of pectin are homogalacturonan, which is composed of a D-galacturonic acid backbone, and rhamnogalacturonan I, which has a backbone consisting of alternating galacturonic acid and rhamnose residues. Both forms have a diverse array of side-chains (3). Pectins are crosslinked with hemicellulose and cellulose and affect plant cell wall pore size, flexibility and strength. Lignin, which adds rigidity to the plant cell wall, is composed of polymers of aromatic residues and is very recalcitrant to deconstruction (4).

Although biochemical activities of select PCWDEs have been investigated in a variety of filamentous fungi, how fungi sense complex carbohydrates in plant biomass, and how that sensing is transduced intracellularly into a hierarchical metabolic response resulting in optimal production of PCWDEs and integration of cellular metabolism is unclear. The production of PCWDEs is dependent on transcription factors that modulate expression of these genes upon appropriate nutrient sensing. In *Neurospora crassa, Aspergillus nidulans, Aspergillus oryzae* and *Penicillium oxalicum,* the transcription factor CLR-2 (ClrB/ManR) is the major regulator of genes involved in the deconstruction of cellulose (5, 6), while in *Trichoderma reesei* and *Aspergillus niger*, the transcription factor Xyr1/XlnR regulates genes involved in both cellulose and hemicellulose degradation (7, 8). In species like *N. crassa* and *Fusarium graminearum*, XlnR homologs regulate genes involved in hemicellulose utilization (9, 10). Transcription factors associated with

pectin deconstruction include RhaR/PDR-1, GaaR and, Ara1. In *A. niger* and *N. crassa,* RhaR/PDR-1 are required for rhamnose utilization (11, 12), while in *B. cinerea* and *A. niger,* GaaR is responsible for galacturonic acid utilization (13, 14). In *A. niger*, the AraR transcription factor modulates arabinose utilization, while a different transcription factor (Ara1) functions in an analogous manner in *Magnaporthe oryzae* and *T. reesei (15, 16)*. Additional transcriptional regulators that affect expression of genes encoding PCWDEs include the carbon catabolite repressor protein CreA/CRE-1 (17, 18), COL-26/BglR (19, 20) and VIB-1/Vib1 (21, 22).

Here we performed transcriptional profiling of *N. crassa* on 40 different carbon sources to provide data on how fungi sense simple to complex carbohydrates and analyzed profiling data to identify regulatory factors associated with carbon source sensing and the regulation of transcriptional responses. From this approach, two transcription factors, one involved in pectin utilization, PDR-2 and one involved in pectin and hemicellulose utilization, ARA-1, were identified and their regulons characterized. Using *in vitro* DNA-affinity purification sequencing (DAP-seq) of transcription factors involved in regulating PCWDE-encoding genes led to a more complete understanding of direct targets of these regulatory proteins and of the crosstalk between transcription factors involved in regulating nutrient sensing on a global level. In particular, our data clarified the role of VIB-1 in the regulation of genes encoding PCDWEs and nutrient scavenging and identified a previously overlooked mechanism of the carbon catabolite repressor protein CreA/CRE-1 in regulating cellular responses to carbon sources.

Results

N. crassa carbon metabolism is distinctly regulated in response to different carbon sources

To improve our understanding of how regulatory networks are integrated during plant biomass utilization by filamentous fungi, we assessed gene expression patterns across 40 different carbon conditions in *N. crassa* (Table S1). To reduce the effects of differential growth on gene expression in different carbon sources, we performed switch experiments where *N. crassa* cells (FGSC2489) were pre-grown in sucrose as a sole carbon source (16 h), washed, and then transferred to media containing the experimental carbon source for 4 h prior to RNA extraction. The

carbon sources were divided into three categories: plant biomass, complex polysaccharides found in the plant cell wall, and the mono- and disaccharide building blocks that make up these complex polysaccharides. We compiled a list of 113 genes encoding predicted PCWDEs in the *N. crassa* genome (Table S2) and assessed expression of this gene set across our carbon panel (Fig. 1A; SI Dataset 1).

At low concentrations, various monosaccharides, disaccharides, and oligosaccharides induce the expression of genes encoding PCWDEs (23, 24). We exposed *N. crassa* to 19 different mono- and disaccharides at 2 mM concentration; this concentration of cellobiose was previously shown to induce robust expression of cellulolytic genes in *N. crassa* (23) (Table S1). As predicted, *N. crassa* induced genes encoding cellulases in response to cellobiose, genes encoding starch-degrading enzymes in response to maltose, genes encoding hemicellulases in response to xylose and arabinose, and genes encoding pectin deconstruction enzymes upon exposure to rhamnose and galacturonic acid (Fig. 1A; SI Dataset 1). However, individual sugars were also capable of inducing expression of PCWDEs not responsible for degrading their parent polymer. For example, cellobiose induced expression of some genes encoding some xylanases and pectinases in addition to cellulases, and arabinose induced expression of some genes encoding some cellulases in addition to arabinases (SI Dataset 1). These data indicate metabolic crosstalk between sugar sensing pathways and/or overlap in regulatory networks. *N. crassa* also showed strong transcriptional responses to complex plant biomass substrates, such as corn stover (a monocotyledonous plant of the grass family) ~~versus~~ and wingnut (*Pterocarya*; a hardwood tree from the walnut family) (Fig. 1A).

Mono-, di-, and oligosaccharides require transport into the cell for utilization and/or signaling for induction of genes encoding PCWDEs. Annotated sugar transporters belong to the Major Facilitator Superfamily (MFS) and led us to hypothesize that uncharacterized sugar transporters would also come from this protein family. To test this hypothesis, we constructed a maximum likelihood tree using protein sequences from all MFS transporters in the *N. crassa* genome (Fig. S1). The majority of predicted sugar transporters with the exception of NCU05897 (fucose permease) and NCU12154 (maltose permease), fell into a single monophyletic clade corresponding to family 2.A.1.1 of the transporter classification database (TCDB; (25)). Of the predicted sugar transporters in this clade, five un-annotated MFS transporters (NCU04537, NCU05350, NCU05585, NCU06384,

NCU07607) had increased expression on unique sugars and complex carbon sources, suggesting potential involvement in catabolism of those carbon sources (Fig. S1; Dataset S1).

To evaluate crosstalk between regulatory pathways that coordinate expression of PCWDEs, we performed weighted gene co-expression network analysis (WGCNA) (26) across the transcriptional dataset and identified twenty-eight modules of co-expressed genes (Fig. 1B; SI Dataset 2) that showed enrichment of specific functional classifications (Fig. S2). The majority of PCWDE genes were found within three modules. Module 1 (red; n= 153) contained gene encoding PCWDEs that are upregulated in response to cellulose and hemicellulose along with notable transcription factors *xlr-1, clr-1, clr-2, hac-1*, and *vib-1* (21, 27, 28). This module also contained 55 genes that encoded hypothetical proteins. Module 2 (yellow; n=42) contained the majority of predicted pectin metabolic genes (28) and eight genes encoding hypothetical proteins. Module 3 (blue; n=42) contained a number of predicted pentose catabolic genes along with some notable xylanases and xylose transporters and nine genes encoding hypothetical proteins (SI Dataset 2; Fig. S2). An additional module (Module 4; n=142; midnight blue) clustered closely with modules 1 and 3. This module was significantly enriched for genes encoding ER and protein processing proteins (cellular transport and protein fate; Fig. S2) that are co-regulated with genes encoding cellulases and xylanases, such as various COPII proteins, SEC-61, KEX2, (SI Dataset 1; S1 Dataset 2). This module also included genes encoding 29 hypothetical proteins.

Defining the PCWDE transcriptional network

Prior studies in *N. crassa* identified conserved transcription factors that are positive regulators of cellulase and some hemicellulase genes (CLR-1/CLR-2), xylanase and xylose utilization genes (XLR-1), pectin-degrading genes (PDR-1), and starch catabolic genes (COL-26) (10, 11, 19, 27). We hypothesized that it would be possible to identify additional regulators involved in plant cell wall degradation by looking for transcription factors with a similar expression profile to a specific class of genes encoding PCWDEs using hierarchical clustering. A systematic analysis of expression profiles of 336 proteins with predicted DNA-binding domains identified 34 additional transcription factors that were specifically induced on different plant biomass components (Table S3A). We hypothesized that strains carrying a deletion

of a transcription factor would display an altered transcriptional profile under the conditions where they were most highly expressed (Table S3B). When the corresponding deletion strains were tested under the respective induction conditions, a majority of the 34 transcription factor deletion mutants did not display a clear expression phenotype as compared to the parental strain, FGSC2489. However, deletion mutants for two transcription factors showed a consistent and obvious role in PCWDE expression, NCU04295 and NCU05414 (SI Dataset 3).

The expression of NCU04295 clustered with genes encoding pectin-degrading enzymes (SI Dataset 3) and a ΔNCU04295 mutant showed decreased expression levels of genes necessary for pectin utilization when grown in presence of pectin-rich citrus peel as compared to wild type cells on citrus peel (Fig. 2A, B; SI Dataset 3). The genes with the largest decrease in expression level in ΔNCU04295 as compared to wild type included pectate lyases genes *ply-1* and *ply-2* (NCU06326 and NCU08176), the galacturonic acid transporter gene *gat-1* (NCU00988), the exo-polygalacturonase genes *gh28-2* (NCU06961), and orthologs of *gaaA*, *gaaB* and *gaaC* (NCU09533, NCU07064 and NCU09532), encoding enzymes for galacturonic acid catabolism (Fig. 2B; SI Dataset 3). The predicted protein sequence of NCU04295 showed similarity (~50% amino acid identity) to GaaR, which plays a role in galacturonic acid metabolism in *B. cinerea* and *A. niger* (13, 14). We therefore named NCU04295 *pdr-2* for pectin degradation regulator-2. Consistent with its predicted function, the Δ*pdr-2* mutant showed a severe growth defect in medium containing pectin or galacturonic acid as the sole carbon source and significantly reduced pectate lyase and endo-polygalacturonanase activity (Fig. 2C,D). A second pectin degradation regulator previously identified in *N. crassa*, *pdr-1*, also shows a severe growth defect on pectin (11). However, unlike Δ*pdr-1* cells, Δ*pdr-2* cells grew on L-rhamnose as the sole carbon source (Fig. S3), suggesting distinct roles for PDR-2 and PDR-1 in regulating pectin degradation. A strain bearing both *pdr-1* and *pdr-2* deletions mimicked the phenotype of either a Δ*pdr-1* or a Δ*pdr-2* mutant (Fig. 2C, D), but did not cause a complete abolition of growth with pectin as the sole carbon source (Fig. S3).

NCU05414 displayed high expression on *Miscanthus* biomass (SI Dataset 1). When compared to wild type cells exposed to 1% *Miscanthus*, a ΔNCU05414 mutant showed reduced expression of genes encoding several arabinosidases (NCU09924, NCU9775), two β-xylosidases (NCU00709, NCU09923), the L-arabinose transporter

lat-1 (NCU02188), and L-arabinitol dehydrogenase *ard-1* (NCU00643) (Fig. 2E; SI Dataset 3), suggesting that the ΔNCU05414 mutant would be defective for utilization of arabinan, arabinose and galactose. As predicted, the ΔNCU05414 strain showed dramatically reduced growth on 2% arabinan, arabinose, and galactose, but was able to metabolize hemicellulose and pectin substrates (Fig. 2F). When NCU05414 was placed under the regulation of the strong constitutive promoter *gpd-1* (oxNCU05414), cells showed increased growth on arabinose relative to wild type (Fig. S3) and increased expression of *ard-1* (LADH; Fig. 2G), further supporting positive regulation of arabinose metabolic genes by NCU05414. The NCU05414 predicted protein showed significant similarity to the Ara1 protein in *T. reesei* and *Magnaporthe oryzae,* where it plays a role in arabinose metabolism and arabinose and galactose catabolism, respectively (16, 29). We therefore named NCU05414 *ara-1*.

Many PCWDEs involved in degradation of heterogeneous substrates like pectin and hemicellulose are under the control of multiple transcription factors. We constructed regulons of transcription factors important for plant biomass deconstruction based on the genes that are differentially expressed between transcription factor mutant versus wild-type cells (*clr-1, clr-2, xlr-1, pdr-2, ara-1*; SI Dataset 3) and from previous studies for *col-26* and *pdr-1* (11, 19). The regulons of CLR-1, CLR-2, XLR-1, PDR-1, PDR-2, ARA-1 and COL-26 showed extensive overlap (Fig. 3). As an example, the expression of the putative acetylxylan esterase gene (*ce1-1* NCU04870), an enzyme responsible for cleaving acetyl groups from xylan and critically important for increasing accessibility of xylan to xylanases, relative to wild type cells showed a 20-fold decrease in expression levels in Δ*clr-2* cells after a shift to Avicel, a 500-fold decrease in expression in Δ*xlr-1* cells after a shift to xylan, and a 7-fold decrease in expression in Δ*pdr-2* cells after a shift to citrus peel (SI Dataset 3). Moreover, the *ce1-1* promoter was shown to be directly bound by both XLR-1 and CLR-2 by chromatin-immunoprecipitation-sequencing (ChIP-seq) (10).

Utilizing DAP-seq to identify direct targets of *N. crassa* transcription factors.

The transcriptional regulons associated with plant biomass deconstruction identified above could be due to direct or indirect regulation of target genes by a particular transcription factor. To define the direct regulons of transcription factors

involved in plant biomass deconstruction, we used DAP-seq, where *in vitro* synthesized transcription factors are used for affinity purification of bound sequences in sheared genomic DNA, which are subsequently identified via DNA sequence analyses (30). To ensure that DAP-seq was an effective method for identifying direct binding sites of transcription factors involved in plant cell wall deconstruction in *N. crassa*, we confirmed the DNA binding sites of CLR-1 and XLR-1, for which ChIP-seq data are available (10).

We re-analyzed promoter regions of genes (defined as within 3 kb of the ATG start site) bound by XLR-1 identified via ChIP-seq (10) (SI Dataset 4) and bound promoter regions identified via DAP-seq data where transcription was reduced by at least $2^{1.5}$ (2.8)-fold via differential RNA-seq analysis of WT versus an Δ*xlr-1* mutant (SI Datasets 3 and 4). We identified 85 XLR-1 target genes using ChIP-seq data and 78 genes via DAP-seq, with 47 genes shared between the two datasets (Fig. S4A,C,F; SI Dataset 4). The binding site sequences from the 78 genes identified in the DAP-seq dataset were used to build an XLR-1 consensus binding motif, which was comparable to the one reported from ChIP-seq data analysis (10) (Fig. S4G). Using the same methods to explore CLR-2, we identified 87 genes with CLR-2-bound promoters via DAP-seq and 65 genes with CLR-2-bound promoters via ChIP-seq; 48 genes were shared between datasets (Fig. S4D, E, F; SI Dataset 3 and 4). Slight differences were identified in the CLR-2 consensus binding sequence using DAP-seq versus that previously reported for ChIP-seq data (10) (Fig. S4G).

Neither the ChIP-seq nor DAP-seq method reliably reduced ~~false positives, which we defined as~~the number of genes whose promoters were bound by the transcription factor, but whose transcription was not differentially expressed between wild type and the transcription factor mutant under the conditions tested. For example, the CLR-2 ~~,~~ ChIP-seq identified 158 genes with promoter regions bound, while DAP-seq identified 1683; however, the majority of DAP-seq bound genes were not differentially expressed in a Δ*clr-2* mutant relative to WT cells. For XLR-1, ChIP-seq identified 1117 genes, while DAP-seq identified 531 (SI Dataset 4). Thus, ~~only through~~ a comparison of ~~with RNA-seq data and~~ differential expression analyses between WT and transcription factor mutants helped to filter~~could~~ the ChIP-seq and DAP-seq datasets ~~be filtered~~ for biologically relevant genes for these specific transcription factors. For the remaining genes whose expression was not altered in the TF mutants, it is unclear whether they are “false positive” or genes

that might be regulated by CLR-2 or XLR-1 under different conditions that were not assessed in this study.

In *T. reesei,* a constitutively active *xyr1* allele (ortholog to *N. crassa xlr-1*) contains a single amino acid substitution (alanine to valine) in the C-terminal predicted alpha helix (31). The construction of the orthologous mutation (A828V) in *N. crassa xlr-1* results in a strain that shows inducer-independent expression and production of hemicellulases (10). To test whether this mutation affected the binding affinity of XLR-1, we also performed DAP-seq on the XLR-1^{A828V} mutant. The binding targets of the XLR-1^{A828V} mutant largely overlapped with the binding targets of XLR-1, indicating that the A828V mutation has little or no influence on XLR-1 DNA binding affinity (Fig. S4B,F; SI Dataset 4).

DAP-seq suggests a multi-tiered system of CRE-1-mediated carbon catabolite repression

CRE-1 is a major regulator of carbon catabolite repression (CCR), a process through which the expression of genes involved in the utilization of non-preferred carbon sources is repressed in the presence of preferred carbon sources (32). Although many PCWDEs are known to be regulated by carbon catabolite repression, it was unclear whether this repression was directly or indirectly mediated by CRE-1. Using DAP-seq, we identified 329 CRE-1 binding sites in 318 promoter regions, with 11 promoters showing two peaks (SI Dataset 4). The 318 genes with promoters bound by CRE-1 were enriched for 30 functional categories (p-value < $1x10^{-5}$) involved in metabolic and catabolic activities (Table S4). The top 17 functional categories were all involved in carbon metabolism, specifically cellulose, hemicellulose, pectin, and starch catabolism, representing approximately 50% of the total CRE-1 peaks and consistent with functions associated with CRE-1. We used the sequences from CRE-1 bound peaks to build a consensus core motif with the best fit core motif being 5'-TSYGGGG-3' (E=2.7x10-23), similar to the 5'-SYGGRG-3' motif described for CreA in *A. nidulans* (33) (Fig. S3C).

If CRE-1 directly represses genes encoding PCWDEs, we would expect to see CRE-1 binding of PCWDE promoter regions. However, only 19 of 113 PCWDE genes had CRE-1 binding sites in the promoter (SI Dataset 4). According to the "double-lock" mechanism proposed for Cre1 in *Aspergillus nidulans* (34), indirect repression of PCWDE expression by CRE-1 could either be due to CRE-1 repression of

transcription factors required for PCWDE gene activation or due to CRE-1 repression of genes necessary to activate those transcription factors. In our DAP-seq dataset, promoters for only two carbon transcription factors were bound by CRE-1, *clr-1* and *ara-1*. However, CRE-1 binding was highly biased for promoters of genes encoding MFS transporters (22 MFS genes), with 15 falling within the major sugar transporter clade (Fig. S1), including one high affinity glucose transporter, *hgt-1* NCU10021 (35) and additional uncharacterized transporters (NCU00809, NCU06522, NCU09287, NCU04537, NCU01494, NCU06384, and NCU05897).

CRE-1 also bound to the promoters of the cellodextrin transporters *cdt-1* (NCU00801), *cdt-2* (NCU08114), and *sut-12/cbt-1* (NCU05853) (36-39). Cells lacking both *cdt-1* and *cdt-2* are unable to activate cellulolytic gene transcription and do not grow on cellulose (40). The binding of CRE-1 to the promoter of *clr-1* likely contributes to the repression of cellulolytic genes by CRE-1, as CLR-1 positively regulates *clr-2*, the major regulator of cellulolytic genes in *N. crassa* (10, 27)(Fig. 4; SI Dataset 4). ~~CRE-1 also bound to the promoters of the cellodextrin transporters *cdt-1* (NCU00801), *cdt-2* (NCU08114), and *sut-12/cbt-1* (NCU05853) (36-39). Cells lacking both *cdt-1* and *cdt-2* are unable to activate cellulolytic gene transcription and do not grow on cellulose (40).~~ Our data therefore suggested that cellulolytic gene expression is repressed by CRE-1 through a combination of direct binding to cellodextrin transporters, the transcription factor *clr-1* and a few cellulolytic PCWDEs ~~, the transcription factor *clr-1*, and cellodextrin transporters~~ (Fig. 4).

For genes involved in hemicellulose deconstruction, CRE-1 binding sites were detected in the promoters of the arabinose-transporter *lat-1* (NCU02188) (28), xylose transporters NCU00821 and NCU04527 (41), the xylodextrin transporter *cdt-2*, which is required for wild type-levels of growth on xylan (38), and pentose transporters *xat-1* (NCU01132) and *xyt-1* (NCU05627) (42) (Fig. 4). CRE-1 binding peaks were not detected in the promoter of the major transcriptional regulator of xylan utilization, *xlr-1*, although CRE-1 binding sites were detected in the promoter of the arabinose utilization regulator, *ara-1*; an Δ*ara-1* mutant showed dramatically reduced growth on arabinan, arabinose, and galactose (Fig. 2). CRE-1 also directly bound to promoters of genes encoding xylanases, galactosidases, and arabinanases, as well as genes necessary for arabinose metabolism (SI Dataset 4). ~~CRE-1 binding peaks were not detected in the promoter of the major transcriptional regulator of xylan utilization, *xlr-1*. However, CRE-1 binding sites were detected in~~

~~the promoter of the arabinose utilization regulator, *ara-1*; an Δ*ara-1* mutant showed dramatically reduced growth on arabinan, arabinose, and galactose (Fig. 2). CRE-1 binding sites were also detected in the promoter of the arabinose transporter *lat-1* (NCU02188) (28), xylose transporters NCU00821 and NCU04527 (41), the xylodextrin transporter *cdt-2*, which is required for wild type levels of growth on xylan (38), and pentose transporters *xat-1* (NCU01132) and *xyt-1* (NCU05627) (42)~~ (Fig. 4).

CRE-1 was not bound to the *pdr-1* or *pdr-2* promoters, which are responsible for regulating the majority of pectinase genes in *N. crassa* (12; Fig. 2A). ~~For genes involved in pectin utilization, DAP-seq showed~~ However, CRE-1 binding sites were identified in the promoter of a major exo-polygalacturonase (NCU06961; *gh28-2*) as well as predicted metabolic enzymes for galacturonic acid utilization (*gaaA* ortholog NCU09533, *gaaB* ortholog NCU07064 and *gaaC* ortholog NCU09532) (Fig. 4; SI Dataset 4). ~~CRE-1 was not bound to the *pdr-1* or *pdr-2* promoters, which are responsible for regulating the majority of pectinase genes. One of the~~An uncharacterized sugar transporter~~s~~ bound by CRE-1, *sut-28* (NCU05897; annotated as a fucose permease; Fig. S1), is a predicted ortholog of the *A. niger* L-rhamnose transporter RhtA (43). The *sut-28* mutant showed reduced growth on L-rhamnose and, to a lesser extent, poly-galacturonic acid (Fig. 5A) and uptake of L-rhamnose in the Δ*sut-28* cells was eliminated (Fig. 5B). ~~Additionally, s~~Similar to a Δ*pdr-1* mutant, Δ*sut-28* cells failed to activate expression of the rhamnose catabolic gene L-rhamnonate dehydratase (NCU09034) (Fig. 5C). The expression of *sut-28* was higher in Δ*cre-1* cells when exposed to L-rhamnose or L-rhamnose and glucose (Fig. S3D) and Δ*cre-1* cells showed increased L-rhamnose uptake as compared to wild type when exposed to pectin and glucose (Fig. 5D). These data support the DAP-seq results indicating that CRE-1 negatively regulates the expression of *sut-28.* ~~Thus, an important component of CRE-1 function includes the repression of transporter genes that play a role in the uptake of signaling molecules acting as inducers of transcription factors and genes associated with cellulose, hemicellulose and pectin utilization (Fig. 4).~~

A Δ*cre-1* mutant shows growth defects relative to its wild-type parental strain when grown on sucrose (Sun and Glass 2011 PLoS One). Previous microarray data of a Δ*cre-1* mutant relative to wild-type under minimal medium conditions with sucrose as the sole carbon source showed that 75 genes showed increased

expression levels (>2 fold) in the Δ*cre-1* mutant (Sun and Glass 2011 PLoS One), seven of which encoded predicted MFS transporters. Of these 75 genes, the promoters of 21 genes were bound by CRE-1 in the DAP-seq dataset, a significant enrichment over random (3.5 genes), and which included all seven of the predicted MFS transporters that showed increased expression in the Δ*cre-1* mutant relative to wild type. These MFS sugar transporters included NCU04537 (monosaccharide transporter), NCU04963 (high affinity glucose transporter), NCU06026 (quinate permease), NCU05897 (*sut-28*), NCU10021 (*hgt-1*), NCU00821(sugar transporter) and NCU05627 (high affinity glucose transporter *ght-1*). The remaining set of 21 genes include a number of carbon metabolic enzymes and 5 genes encoding proteins of unknown function (SI Dataset 4). Thus, an important component of CRE-1 function includes the repression of transporter genes that play a role in the uptake of signaling molecules acting as inducers of transcription factors and genes associated with cellulose, hemicellulose and pectin utilization (Fig. 4).

~~Thus, an important component of CRE-1 function includes the repression of transporter genes that play a role in the uptake of signaling molecules acting as inducers of transcription factors and genes associated with cellulose, hemicellulose and pectin utilization (Fig. 4).~~

DAP-seq of VIB-1 reveals a global role in regulating carbon metabolism

VIB-1 is a Zn_2Cys_6 transcription factor first identified for its role in mediating self/nonself recognition and heterokaryon incompatibility in *N. crassa* (44, 45). The Δ*vib-1* mutant also shows severely reduced growth on Avicel and a weak induction of *clr-2* (21), a phenotype also observed in *T. reesei* Δ*vib1* strains (22). In addition to Avicel, the Δ*vib-1* mutant also has a severe growth defect on pectin and a moderate growth defect on xylan (Fig. S5A).

RNA-seq was previously performed on Δ*vib-1* cells exposed to Avicel and carbon starvation conditions (21). Here, we performed additional RNA-seq experiments on the Δ*vib*-1 mutant exposed to 1% pectin or 1% xylan as the sole carbon source, 1% BSA as the sole carbon and nitrogen source, and 1% ground *Miscanthus* as the complete nutrient source (SI Dataset 5). RNA-seq data reflected the severity of growth phenotypes, as exposure to Avicel, pectin, and BSA displayed the greatest number of differentially expressed genes between WT and the Δ*vib-1*

mutant. Consistent with its phenotype, the Δ*vib-1* mutant has a more similar expression profile to WT cells under xylan conditions.

Using DAP-seq, we identified VIB-1 binding sites within 1.5 kb upstream of the ATG start site of 1,742 genes (SI Dataset 4). The RNA-seq datasets were utilized to eliminate false positives from DAP-seq data by limiting the set to genes with at least a $2^{1.5}$ (2.8)-fold change in gene expression in any of our six conditions. In total, we identified 238 direct target genes of VIB-1 (SI Dataset 5). Hierarchical clustering of gene expression data of these direct targets showed that one cluster included the majority of genes that were down regulated in the Δ*vib-1* mutant in more than three conditions. We considered these genes to be the core regulon of VIB-1 (Fig. 6A). A consensus binding motif from VIB-1 peaks within the 1.5 kb promoter regions of core regulon genes showed conservation of three critical bases: T, A and C (Fig. 6C~~B~~).

The 56 gene VIB-1 core regulon included genes involved in heterokaryon incompatibility (*tol, pin-c* and *het-6*) and a number of uncharacterized genes encoding proteins with predicted roles in heterokaryon incompatibility (HET domain proteins and genes with polymorphic alleles in wild populations; NCU03533, NCU05840, NCU07335 and NCU04453) (SI Dataset 5) (46). Most of the other annotated genes in the VIB-1 core regulon were associated with metabolism, including three arabinofuranosidases (NCU09170 NCU09975 and NCU02343), a beta-xylosidase (NCU09923), three cellulose PMOs (NCU02240, NCU09764 and NCU02344), a starch active PMO (NCU08746), a galacturonic acid transporter (*gat-1*; NCU00988), an exogalacturonase (NCU06961), rhamnogalacturonan acetylesterase (NCU09976), a secreted phospholipase (NCU06650), and acid phosphatase (*pho-3;* NCU08643) (SI Dataset 5) (Fig. 6B). Promoters of three genes encoding LaeA-like methyltransferase domains (NCU05841, NCU05832, and NCU05501) were in the core VIB-1 regulon and four additional LaeA-like genes were direct targets of VIB-1 (NCU04909, NCU04717, NCU04707, NCU01148) (SI Dataset 5). LaeA is a regulator of secondary metabolism in ascomycete fungi first described in *A. nidulans (47)*.

The *clr-2* and *pdr-2* genes were the only ones encoding transcription factors that were direct targets of VIB-1 (Fig. 6B). In the Δ*vib-1* mutant, expression of *clr-2* was reduced 5.2-fold relative to wild type during exposure to Avicel and expression of *pdr-2* was reduced 3.4-fold relative to wild type during exposure to pectin. In addition to *clr-2* and *pdr-2,* a number of PCWDE-encoding genes were bound and regulated by VIB-1, including genes encoding enzymes in the core VIB-1 regulon (above), cellulases (*gh6-3*, NCU07190; *gh45-1*, NCU05121, NCU05751), arabinosidase (NCU05965), rhamnogalacturonase (NCU05598), rhamnogalacturonan acetylesterase (NCU09976), a pectinesterase (NCU10045), xylanases (NCU02855, NCU04997), feruloyl esterase B (NCU09491), and acetyl xylan esterases (NCU08785, NCU04494) (Fig. 6C). Additional genes encoding PCWDEs that were down-regulated in the Δ*vib-1* mutant, but that did not have VIB-1 binding sites in their promoters, could be explained by reduced expression of *clr-2* or *pdr-2* (Fig. 6C), consistent with the severe growth defect on cellulose and pectin substrates in the Δ*vib-1* mutant.

Our DAP-seq data suggests that VIB-1 acts through *clr-2* to promote cellulase gene expression. However, ChIP-seq identified *vib-1* as a target of the cellulase regulator, CLR-1 (10). CLR-1 also binds to the promoter and is required for the expression of *clr-2 (10).* These observations suggest an interplay in the regulation of *clr-2* by CLR-1 and VIB-1. To investigate these interactions, we measured cellulase production in a Δ*vib-1* Δ*clr-3* strain, where repression of CLR-1 activation in the absence of cellulose is relieved (48) and in a Δ*vib-1* Δ*cre-1* mutant, which eliminates regulation of *clr-1* by CRE-1. Both double mutant strains showed higher cellulase activity than Δ*vib-1* cells (p-adj<0.01), indicating that when relieved from either CLR-3-or CRE-1-mediated repression, CLR-1 was capable of activating cellulolytic gene expression in the absence of VIB-1 (Fig. S5B). However, the cellulase activity of Δ*vib-1* Δ*clr-3* or Δ*vib-1* Δ*cre-1* cells was not as high as wild type cells, indicating that CLR-1 and VIB-1 were both required for full activation of cellulase genes in *N. crassa* (p-adj<0.01) (Fig. S5B).

In addition to defects in growth on cellulose and pectin, *N. crassa* and *A. nidulans vib-1/xprG* mutants show reduced growth when BSA is the sole carbon or nitrogen source (44, 49). However, analyses of the VIB-1 regulon on BSA did not reveal a clear reason for this growth deficit. Only three genes encoding predicted proteases/peptidases were significantly reduced in expression level in the Δ*vib-1*

mutant as compared to wild type cells, including a metalloprotease (*mpr-8*; NCU07200), a carboxypeptidase (*mpr-14;* NCU07536) and a proteinase T (*spr-7;* NCU07159). An additional set of vitamin B6 synthesis genes also showed decreased expression in the Δ*vib-1* mutant specifically on BSA, including *pdx-1* (NCU06550) and *pdx-2* (NCU06549) that encode proteins that form the enzyme complex pyridoxal 5'- phosphate synthase or vitamin B6 synthase (SI Dataset 5). Pyridoxal 5'-phosphate is a cofactor for many enzymes involved in amino acid metabolism and other protein metabolic processes (50).

DISCUSSION

In nature, the primary source of nutrients for *N. crassa* is plant biomass. In this study, we determined expression patterns of the laboratory strain of *N. crassa* to different types of carbon sources, including mono-, di-, oligosaccharides and plant biomass. These results showed that *N. crassa* responds specifically to the constituents of plant biomass in a largely specific manner (e.g. genes encoding cellulases were induced upon exposure of *N. crassa* to cellobiose), but also revealed cross regulation of genes encoding enzymes not found in the substrate (e.g. genes encoding some xylanases were induced upon exposure of *N. crassa* to cellobiose). Induction of PCWDEs by constituents of the plant cell wall, particularly cellobiose and xylose, have also been shown for other basidiomycete and ascomycete fungi (51-53). These data indicate that filamentous fungi respond specifically to the presence of the individual nutrient sources available, but also that the cells anticipate the presence of additional nutrient sources. This anticipation is likely due to the fact that individual components of the plant cell wall are unlikely to be found alone in nature and therefore expression profiles of fungi deconstructing plant biomass are shaped by the structure and composition of the plant cell wall.

Analyses of a large dataset of microarray transcriptomics data of *A. niger* exposed to different conditions and performed by multiple laboratories was used to generate co-expression networks (54). Here, WGCNA on *N. crassa* datasets from exposure to different carbon sources under carefully controlled conditions identified 28 clusters of co-regulated genes (SI Dataset 2). We were particularly interested in defining new transcription factors and regulons associated with plant biomass deconstruction and identified 34 transcription factors whose expression level varied across our panel. Of these, two transcription factor mutants, Δ*ara-1* and Δ*pdr-2*

showed a significantly different response to *Miscanthus* and pectin, respectively, as compared to WT cells (Fig. 2) and a deficiency in the utilization of arabinose/galactose (Δ*ara-1*) and galacturonic acid and pectin (Δ*pdr-2*). Our transcriptional analyses showed that the expression of the *lat-1* transporter gene and the *ard-1* gene were significantly down-regulated in the Δ*ara-1* mutant. Loss of LAT-1 prevents arabinose transport (55), while *ard-1* encodes L-arabinitol-4-dehydrogenase, which catalyzes the second reaction of arabinose catabolism (56) as well as the third step of the oxidoreductive galactose catabolism (57). PDR-2 is involved in the regulation of genes encoding homogalacturonan backbone-degrading enzymes and galacturonic acid catabolic enzymes, similar to GaaR in *A. niger* and *B. cinerea* (13, 14). Activation of a number of pectinase genes, such as the endo-PGase *gh28-1*, were dependent on the presence of both PDR-1 and PDR-2 (Fig. 2B,D). Further characterization of transcription factors associated with plant biomass deconstruction, including those identified in this study, will lead to a better understanding of metabolic crosstalk and reveal direct and/or indirect influence on each other in a synergistic regulatory network important for temporal and spatial deconstruction of plant biomass.

To define the direct regulons of transcription factors involved in plant biomass deconstruction, we utilized DAP-seq, developed to assess the direct targets of predicted transcription factors in *Arabidopsis thaliana* (30). Unlike other methods of identifying DNA binding sites, DAP-seq has the advantage that chromatin structure and growth conditions do not play a role in determining transcription factor binding sites. However, transcription factors that require chromatin structure or other co-factors to bind to their DNA target site will not be identified by DAP-seq. Our comparison of ChIP-seq and DAP-seq data for CLR-2 and XLR-1 showed a strong overlap in these two datasets. Analyses of both datasets were helped substantially by the availability of RNA-seq data under different carbon sources and by profiling data of the transcription factor mutants under these same conditions. ~~We may have missed direct targets of transcription factors using either DAP-seq/RNAseq or ChIPseq/RNAseq methods due to our stringent differential expression requirements (at least $2^{1.5}$-fold) from expression analyses taken at a single time point.~~ Although we identified 34 TFs whose expression varied across our transcriptional profiling dataset, mutants in a majority of these TFs did not show an obvious expression profile difference to wild type when shifted to conditions where their expression

increased. This result could be due to redundancy of TF function in nutrient regulation, a role of the TF at a different time point that what was assessed in this study, or a role in cross regulation that was not obvious from the RNA-seq dataset. We predict that these TFs do play a role in nutrient regulation in *N. crassa* and that a combination of DAP-seq to help identify conditions for RNA-seq studies and expression profiling at additional time points may help to illuminate their function. We may also have missed direct targets of transcription factors using either DAP-seq/RNAseq or ChIPseq/RNAseq methods due to our stringent differential expression requirements (at least $2^{1.5}$-fold) from expression analyses taken at a single time point.**Nonetheless,** ~~this approach was particularly helpful in defining the role of VIB-1 and elucidating additional functions of CRE-1.~~

Our data shows that CRE-1-mediated carbon catabolite repression acts not only through regulation of PCWDEs and their positive transcription factor regulators, but also through key sugar catabolic genes and sugar transporters. Repression of transporter gene expression by CRE-1 reduces entry of signal-transducing sugars into the cell, thus limiting induction of genes encoding PCWDEs. In *A. niger*, low concentrations of galacturonic acid were required to induce gene expression of galacturonic acid utilization genes, including a galacturonic acid transporter, which was repressed by glucose in a CreA-dependent manner (58). Thus, CRE-1 may be regulating CCR through more than 4 levels of control or a “quadruple lock” mechanism: 1. Regulating expression of sugar transporters, 2. Regulating expression of sugar catabolic genes, 3. Regulating expression of transcription factors important for expression of genes encoding PCWDEs, and 4. Regulating the expression of genes encoding PCWDEs (Fig. 4). This “quadruple lock” mechanism may be important in nutrient sensing, production of PCWDEs based on nutrient source, and for integration of different nutrient signals for optimal metabolic regulation during plant biomass deconstruction. Our DAP-seq data on CRE-1 provides a framework for investigating the variety of conditions where CRE-1 plays a role in regulating metabolism, particularly in conjunction with transcription factors that control condition-specific responses.

The transcription factor VIB-1 belongs to the p53 superfamily, which in mammalian cells regulates the cell cycle, DNA repair, and apoptosis (59). In *Saccharomyces cerevisiae*, the p53 homolog, Ndt80, regulates entry into meiosis upon nitrogen starvation (60). The genome of *N. crassa* has three p53 homologs,

612 *vib-1*, *fsd-1* and NCU04729, none of which are required for meiosis, although both
613 *fsd-1* and *vib-1* mutants affect female reproductive structure development, which is
614 regulated by nutritional status (61). In other filamentous fungi, *vib-1* homologs have
615 been shown to regulate protease production, production of extracellular hydrolases
616 and PCWDEs, N-acetyl glucosamine catabolism, and secondary metabolism (62).
617 Additionally, a *vib-1* homolog in the human pathogen *Candida albicans* regulates
618 virulence (63). These observations suggest a general role for VIB-1 orthologs in
619 sensing and responding to the availability of nutrients in their environment. Upon
620 starvation, VIB-1 is required for an increase in the expression of a number of
621 secreted proteins associated with polysaccharide and protein degradation (VIB-1
622 core regulon). These “scout” enzymes release mono-/di-/oligosaccharides, which are
623 transported into the cell, resulting in full activation of genes and secretion of
624 enzymes associated with the utilization of a particular plant biomass component.
625 This model is consistent with VIB-1 functioning as a general starvation response
626 transcription factor, or a transcription factor important for basal expression of
627 nutrient acquisition genes. In cells lacking VIB-1, a positive feedback loop is not fully
628 initiated, and full expression PCWDE genes necessary for optimal utilization of plant
629 biomass is not achieved. Also consistent with this model, is that *vib-1* is not under
630 carbon catabolite repression regulation, as VIB-1 and its direct target genes are not
631 regulated by CRE-1.
632 Previously, it was hypothesized that VIB-1 functions upstream of CRE-1 and
633 COL-26, as the introduction of Δ*cre-1* Δ*col-26* mutations into a Δ*vib-1* mutant
634 suppressed the inability of the Δ*vib-1* mutant to utilize cellulose (21). Our DAP-seq
635 and RNA-seq data supports an alternative hypothesis. We predict that the deletion
636 of *cre-1* and *col-26* allows sufficient expression of *clr-2* to restore growth of the Δ*vib-*
637 *1* mutant on cellulose (partly due to lack of repression of the cellodextrin
638 transporters by CRE-1) in a manner similar to how deletion of the three β-
639 glucosidase genes (Δ3BG) restored cellulase production in the Δ*vib-1* mutant on
640 cellobiose (48). Under carbon-limiting conditions, VIB-1 promotes expression of *clr-2*
641 and *pdr-2* along with a small set of PCWDEs. These secreted enzymes cleave plant
642 biomass and signaling sugars are transported into the cell. For cellulose utilization,
643 cellobiose (or a modified version of cellobiose) results in inactivation of the
644 repressor CLR-3 (48), allowing activation of CLR-1. CLR-1 promotes expression of
645 *clr-2* and cellulases, and together with VIB-1, results in full expression of *clr-2* and

induction of a positive feedback loop. As the glucose concentration increases inside the cell, CRE-1 mediated CCR is activated, reducing expression of *clr-1* and cellodextrin transporters *cdt-1, cdt-2*, and *cbt-1,* thus negatively regulating expression of PCWDEs both by limiting the expression of *clr-1* and the cleaving and import of sugar signaling molecules (Fig. 4). Our data supports the cooperative regulation of PCWDEs by negative regulation of transporters by CRE-1 and positive regulation of enzyme scouts that regulate signaling processes via VIB-1.

VIB-1 regulation of HET domain genes may also play a role in nutrient acquisition. HET domain genes allow fungi to distinguish between self and non-self cells, and initiate programmed cell death upon fusion of non-self pairs (64). Starvation increases vegetative cell fusion frequency in a number of ascomycete fungi, including *N. crassa* (65-67). We hypothesize that VIB-1 increases expression of these HET domain genes to ensure viable fusion is prevented between non-self cells. Potentially, this activity may also be related to the regulation of secondary metabolism by VIB-1-like proteins. The promoters of LaeA-like methyltransferase domain containing proteins were abundant in the direct target gene set of VIB-1. LaeA and LaeA-like methyltranferase orthologs are negative regulators of secondary metabolite production in fungi (47, 68, 69). The modulation of the expression of these methyltransferases by VIB-1 may have downstream gene regulatory consequences that may affect competition among microbes and nutrient acquisition during plant biomass deconstruction and utilization.

Materials and Methods

Comprehensive list of PCWDE genes in the *N. crassa* genome.

A comprehensive list of predicted *N. crassa* genes encoding PCWDE was compiled by examining all CAZymes from the Carbohydrate Active Enzymes Database (http://www.cazy.org) (70) (Table S2).

Strains, Growth Conditions, RNA extraction and RNA-seq

Strains are listed in Table S5 (see Supplemental material and Methods for strain construction). For RNAseq experiments induction conditions, 2mM mono and disaccharides were used (23); for complex polysaccharides and plant biomass 1% (w/v) was used (Table S1). RNA isolation and RNA-seq methods are ~~as~~ described in ~~(~~Supplemental Materials and Methods~~)~~. Filtered reads were mapped against *N.*

crassa OR74A genome (v12) using Tophat 2.0.4 (71) and transcript abundance was estimated with Cufflinks 2.0.2 (72) in fragments per kilobase of transcript per million mapped reads (FPKMs) using upper quartile normalization. Differential expression analysis was performed on raw counts with DEseq2 version 3.3 (Love MI, Huber W, & Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550) using data from biological triplicates. Data available at: https://genome.jgi.doe.gov/portal/TheFunENCproject/TheFunENCproject.info.html. https://www.ncbi.nlm.nih.gov/sra submission in progress.

Weighted Correlation Network Analysis (WGCNA) and FUNCAT analyses

The gene co-expression network was calculated across expression profiles for the wild-type strain exposed to carbon sources listed in Dataset S1 using the R package WGCNA (26) and a custom catalogue (11) based on MPS Functional Catalogue Database (FuncatDB) (73) with expanded categories for cell wall degradation-related genes for enrichment analysis.

Enzyme activity and transport assays

WT and Δ*cre-1* strains were induced in 0.5% pectin or 0.5% pectin plus 2% D-glucose and transferred to either 100 µM L-rhamnose or 100 µM L-rhamnose plus 100 µM D-glucose as uptake solution. WT and Δ*sut-28* strains were transferred to uptake solutions containing either 100 µM L-rhamnose, 90 µM D-fucose (VWR, A16789), 90 µM D-xylose or 90 µM D-galactose (Sigma Aldrich, G0750). Monosaccharide concentration of sample supernatants was quantified by high pH Anion Exchange Chromatography – Pulsed Amperometric Detection (HPAEC-PAD) on an ICS-3000 instrument (Thermo Scientific, USA). A 25 µl sample was injected onto a Dionex CarboPac PA20 column (3 × 30 mm guard and 3 × 150 mm analytical) and eluted using an isocratic mobile phase of 10 mM NaOH at 0.4 ml/min and 30°C over 12 min. Cellulase activity assays were modified from Coradetti *et al* (27) (see Supplemental Material and Methods).

DAP-seq

Predicted open reading frames for each transcription factor were amplified from cDNA generated using RNA to cDNA EcoDry premix (Clonetech). Amplified transcription factor sequences were inserted into an expression vector containing T7 and SP6 promoters upstream of HALO tag as previously described (30).

In vitro transcription and translation of transcription factors was performed using Promega TnT T7 Rabbit Reticulocyte Quick Coupled Transcription/Translation System by incubating 1ug of plasmid DNA with 60ul of TnT Master Mix and 1.5ul of 1mM methionine overnight at room temperature. Expression was verified using Western blot analysis with Promega Anti-HaloTag monoclonal antibody. Single DAP-seq libraries were generated once for each transcription factor tested and sequenced once with Illumina MiSeq 2x150BP runs.

Filtered reads were aligned to *N. crassa* OR74A genome (v12) using Bowtie2 v2.3.2 (71). Peak calling was performed using MACS2 v2.1.1 (74) with p-value cutoff at 0.001 and utilizing negative control library alignments. Peaks within 3000 bp upstream of translation start sites were selected for and annotated using a custom python script. The same python script was used for reanalysis of ChIP-seq peaks dataset from Craig et al. 2015 (10) for DAP-seq/ChIPseq comparisons. DAP-seq data available at: https://www.ncbi.nlm.nih.gov/sra/SRP133627.

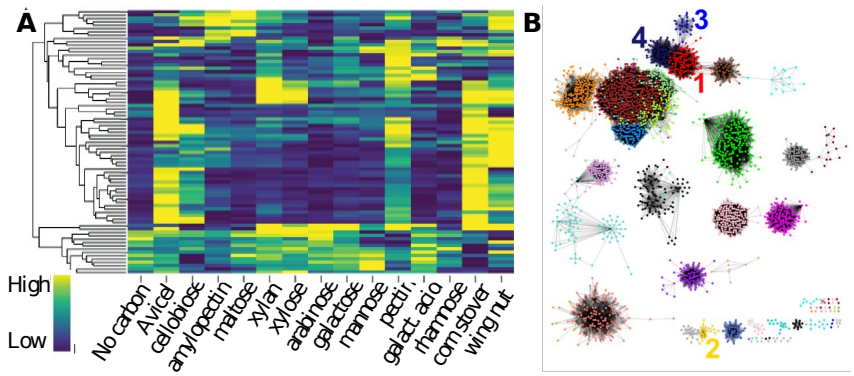
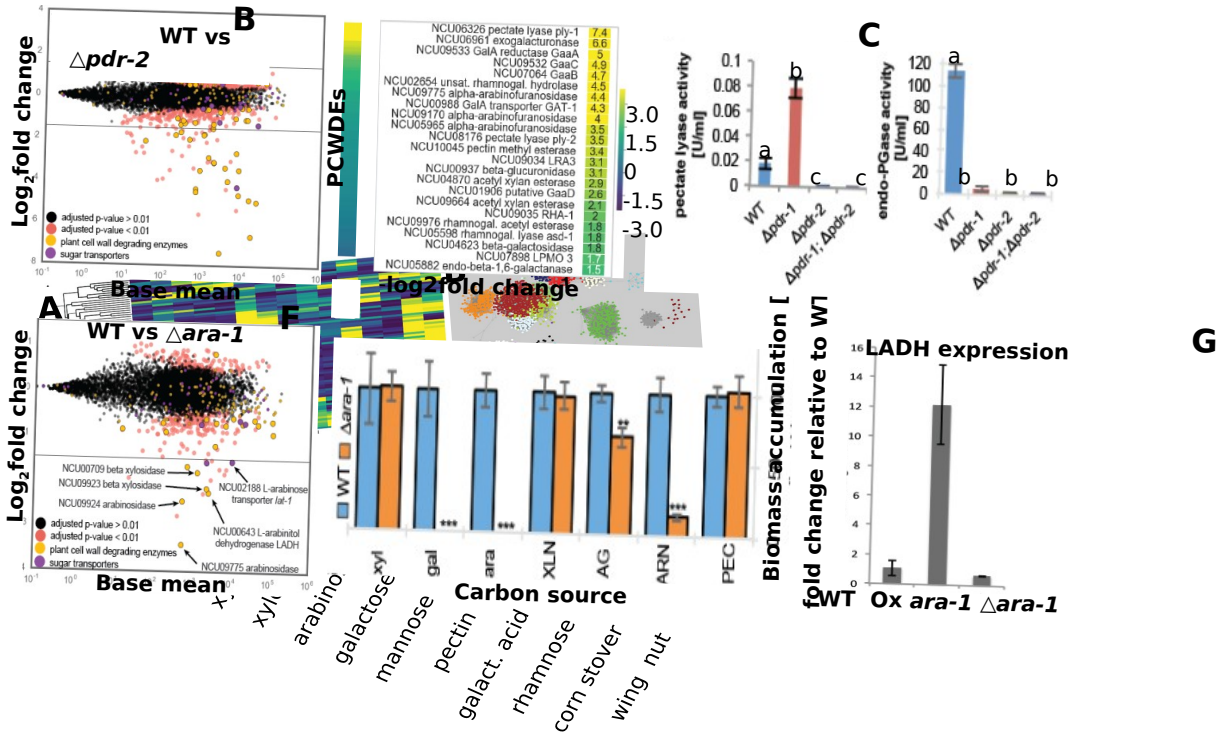


Figure 1. Hierarchical clustering and weighted gene co-expression network analysis of *N. crassa* transcriptome across carbon sources. (A) Hierarchical clustering of of the normalized counts (FPKMs) ~~expression patterns~~ of genes encoding PCWDEs in cells shifted to indicated carbon sources. All di- and monosaccharides are at 2mM concentration and complex carbohydrates are 1% W/V. Color bar represents the spectrum from lowest normalized count to highest normalized count for each gene centered on mean expression. (B) Co-expression network with nodes representing genes colored by modules and edges between genes with correlated expression profiles, shown using Cytoscape (75) (SI Dataset 2). Four modules enriched in genes encoding PCWDEs and polysaccharide metabolism are labeled. Module 1 (red): genes associated cellulose and hemicellulose utilization. Module 2 (yellow): genes associated with pectin deconstruction. Module 3 (blue): pentose catabolic and xylan utilization genes. Module 4 (midnight blue): genes encoding ER and protein processing proteins (SI Dataset 2). Total number of genes shown in the network is 3,282.

758


A
B
WT vs
△pdr-2
Log2fold change
Base mean
adjusted p-value > 0.01
adjusted p-value < 0.01
plant cell wall degrading enzymes
sugar transporters
PCWDEs
NCU06326 pectate lyase ply-1
NCU06961 exogalacturonase
NCU09533 GalA reductase GaaA
NCU09532 GaaC
NCU07064 GaaB
NCU02654 unsat. rhamnogal. hydrolase
NCU09775 alpha-arabinofuranosidase
NCU00988 GalA transporter GAT-1
NCU09170 alpha-arabinofuranosidase
NCU05965 alpha-arabinofuranosidase
NCU08176 pectate lyase ply-2
NCU10045 pectin methyl esterase
NCU09034 LRA3
NCU00937 beta-glucuronidase
NCU04870 acetyl xylan esterase
NCU01906 putative GaaD
NCU09664 acetyl xylan esterase
NCU09035 RHA-1
NCU09976 rhamnogal. acetyl esterase
NCU05598 rhamnogal. lyase asd-1
NCU04623 beta-galactosidase
NCU07898 LPMO 3
NCU05882 endo-beta-1,6-galactanase
3.0
1.5
0
-1.5
-3.0
log2fold change
C
pectate lyase activity [U/ml]
endo-PGase activity [U/ml]
WT
Δpdr-1
Δpdr-2
Δpdr-1;Δpdr-2
D
WT vs △ara-1
NCU00709 beta xylosidase
NCU09923 beta xylosidase
NCU09924 arabinosidase
NCU02188 L-arabinose transporter lat-1
NCU00643 L-arabinitol dehydrogenase LADH
NCU09775 arabinosidase
F
WT
Δara-1
Carbon source
xyl
arabino
galactose
mannose
pectin
galact. acid
rhamnose
corn stover
wing nut
XLN
AG
ARN
PEC
Biomass accumulation [
G
LADH expression
fold change relative to WT
WT
Ox ara-1
△ara-1

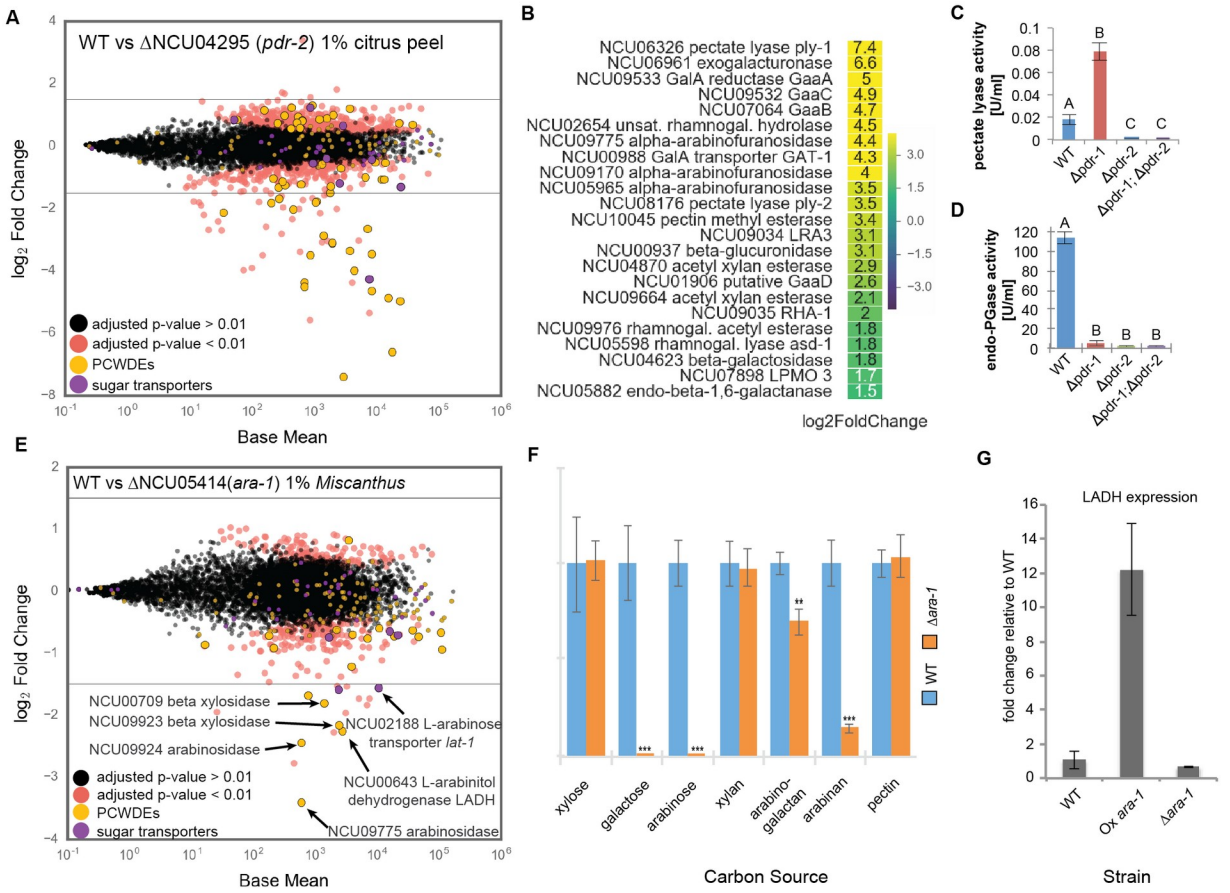


Figure 2. The transcription factor *pdr-2* regulates pectin degradation and the transcription factor *ara-1* regulates arabinose utilization. (A) Differential expression analysis of ΔNCU04295 (*pdr-2*) relative to wild type cells after a shift to 1% w/v citrus peel (SI Dataset 1). PCWDEs are yellow and sugar transporters are purple. "Base mean" is the mean of normalized counts for triplicates of both conditions tested. (B) Differential expression of PCWDEs ranked by degree of $\log_2$ fold change from (A). (C) pectate lyase and (D) endo-polygalacturonanase (endo-PGase) activities of Δ*pdr-1*, Δ*pdr-2* and Δ*pdr-1* Δ*pdr-2* mutants relative to WT. Error bars represent standard deviation (n = 3). Significance was determined by ANOVA followed by a post-hoc Tukey's test. The letters above each bar indicate statistical significance with a mean difference of $p < 0.05$. (E) Differential expression analysis of ΔNCU05414 (*ara-1*) in comparison to WT after cultures were shifted to 1% w/v Miscanthus (SI Dataset 1). PCWDEs are in yellow and sugar transporters are in purple. "Base mean" is the mean of normalized counts for triplicates of both conditions tested. (F) Relative biomass accumulation of Δ*ara-1* normalized to WT cultured in the indicated carbon sources. Significance was determined by an

independent two-sample t-test of WT against Δ*ara-1* with **p < 0.01 and ***p < 0.001 (n=3). (G) Relative *ard-1* (L-arabinitol dehydrogenase: LADH) expression relative to *act* after shift to arabinose in Δ*ara-1* and in an *ara-1* overexpression strain (Ox *ara-1*). For D, F and G, error bars represent standard deviation (n >=3).

.

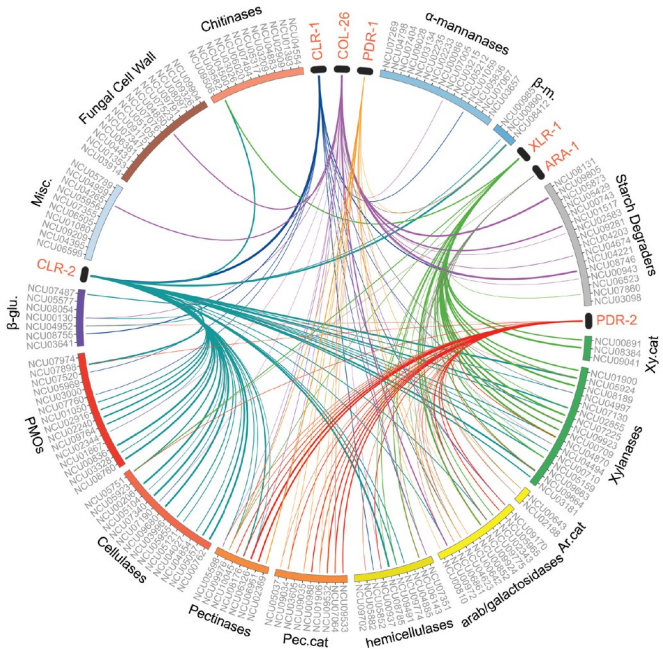


Figure 3. Overlapping regulons of major PCWDE regulators in *N. crassa*. Plot built with Circos v0.69 (76) to display positive regulation of catabolic CAZymes by indicated transcription factors (red) (SI Dataset 1). RNA-seq data for all TFs was obtained under identical culture conditions (see materials and methods). CAZymes are divided into functional groups displayed on the outer edge of the plot. Pec. cat" pectin catabolism, "Ar" arabinose, "Ar. cat." Arabinose catabolism, "Xy" xylose and "β-m" for β–mannanases, β-glu for β-glucosidases, PMO, polysaccharide monooxygenases. Each CAZyme is represented by its gene ID. Each line represents genes with significantly different expression between WT and a transcription factor deletion mutant under the following conditions: Δ*clr-1* and Δ*clr-2* shifted to 1% Avicel, Δ*col-26* shifted to 2mM maltose (Yi), Δ*pdr-1* shifted to 1% pectin (Thieme), Δ*pdr-2* shifted to 1% citrus peel, Δ*ara-1* shifted to 1% *Miscanthus*, and Δ*xlr-1* shifted to 1% xylan (SI Dataset 3). The thickness of the line corresponds to degree of fold change in the transcription factor deletion mutants as compared to wild type cells (11, 19) ~~(SI Dataset 3)~~.

797

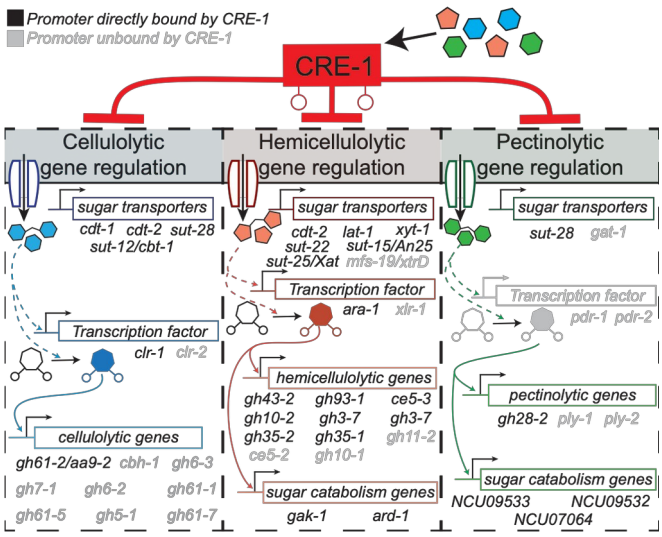


Figure 4. CRE-1-mediated carbon catabolite repression acts through sugar transporter, transcription factor, sugar catabolism, and PCWDE genes to regulate plant cell wall degradation. CRE-1 regulates expression of PCWDE regulons by repressing expression of sugar transporters, transcription factors, and genes involved in the utilization of plant biomass components. Sugars transported into the cell may play either a direct or indirect role in the activation of transcription factors necessary for cellulolytic, hemicellulolytic, and pectinolytic gene expression. The promoters of genes in black are directly bound by CRE-1, and the promoters of genes in grey are not bound by CRE-1. Blue, orange, and green arrows indicate regulation that occurs downstream of CRE-1-mediated repression.

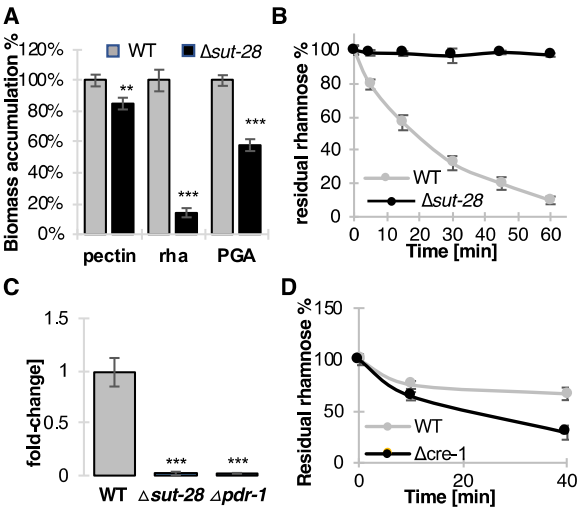


Figure 5. *sut-28* expression and rhamnose transport activity in wild type and Δ*cre-1* strains. (A) Relative biomass of FGSC2489 (WT) and Δ*sut-28* strains incubated in pectin, rhamnose, and polygalacturonic acid as determined by dry weight. (B) Rhamnose uptake in FGSC2489 and Δ*sut-28* strains after induction on pectin. (C) Relative NCU09034 (L-rhamnonate dehydratase) expression relative to *act* in FGSC2489, Δ*sut-28*, and Δ*pdr-1* after induction on rhamnose. (D) Rhamnose uptake in FGSC2489 and Δ*cre-1* strains induced with pectin plus glucose. Error bars represent standard deviation (n >=3). Significance was determined by an independent two-sample t-test of WT against Δ*ara-1* with **p < 0.01 and ***p < 0.001.

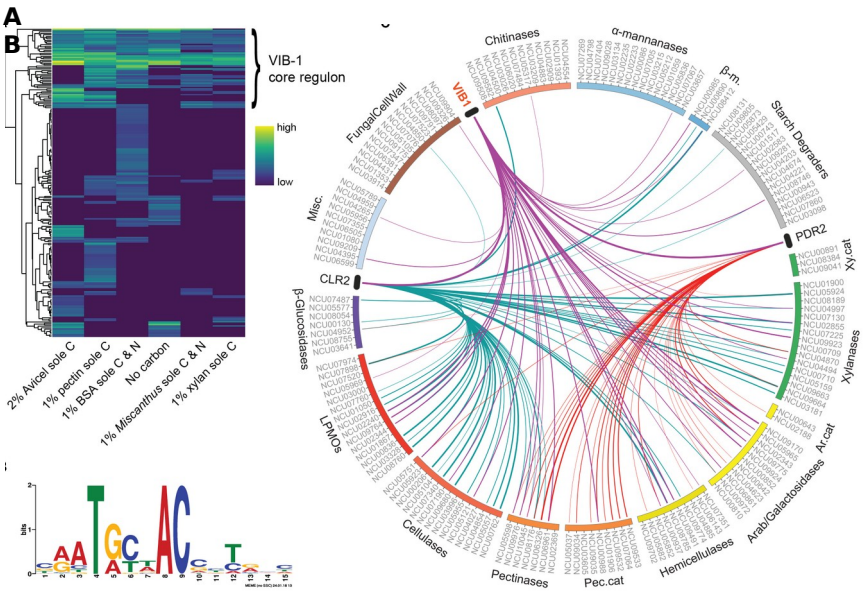


Figure 6. VIB-1 regulon. (A) Hierarchical clustering of $\log_2$ fold change values from differential expression analysis of FGSC2489 versus Δ*vib-1* strains shifted to the indicated carbon conditions. Only genes with greater than $2^{1.5}$-fold change and with promoters bound by VIB-1 via DAP-seq are included. The VIB-1 core regulon is a cluster of genes that were differentially expressed across multiple conditions (SI Dataset 5). (B) VIB-1 binding motif built using MEME v4.12.0 using DAP binding peak sequences of VIB-1 core regulon. E-value = 1.8^{-89}. (C) Plot built with Circos v0.69 (76) to display positive regulation of a catabolic CAZymes by VIB-1 and the transcription factors CLR-1 and PDR-2, which are bound and directly regulated by VIB-1. The thickness of the line corresponds to degree of fold change between WT and transcription factor mutant (SI Dataset 3 and 5).

Acknowledgements

We acknowledge the use of deletion strains generated by grant P01 GM-068087 "Functional Analysis of a Model Filamentous Fungus," and which are publicly available at the Fungal Genetics Stock Center. This work was supported by an Energy Biosciences Institute Grant, ~~to N.L.G. and~~ a Laboratory Directed Research and Development Program of Lawrence Berkeley National Laboratory under US Department of Energy Contract DE-AC02-05CH11231 and funds from the Fred E. Dickinson Chair of Wood Science and Technology to N.L.G. ~~.~~ V.W.W was partially supported by a National Institutes

of Health NRSA Trainee Grant 5T32GM007127-39. We thank ~~the Fungal Genetics Stock Center for strains and~~ Elias Bleek (TUMS) for assistance with transporter assays.

V.W.W., N.T., L.B.H., A.D., D.J.K., J.L., V.S., A.L., R.M. and M.J.B. performed experiments. V.W.W., S.C. and Y.X. performed computational analyses. V.W.W., J.P.B., R.C.O., I.V.G. and N.L.G. were involved in study conception and design. V.W.W., N.T., L.B.H., J.P.B. and N.L.G. wrote and edited the manuscript.

References **NOT UPDATED**

1. Fernandez CW & Kennedy PG (2016) Revisiting the 'Gadgil effect': do interguild fungal interactions control carbon cycling in forest soils? *New Phytol* 209:1382-1394.
2. Gupta VK*, et al.* (2016) Fungal enzymes for bio-products from sustainable and waste biomass. *Trends Biochem Sci* 41:633-645.
3. Caffall KH & Mohnen D (2009) The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. *Carbohyd Res* 344:1879-1900.
4. Ragauskas AJ*, et al.* (2014) Lignin valorization: improving lignin processing in the biorefinery. *Science* 344:1246843.
5. Kunitake E & Kobayashi T (2017) Conservation and diversity of the regulators of cellulolytic enzyme genes in Ascomycete fungi. *Curr Genet* 63:951-958.
6. Huberman LB, Liu J, Qin L, & Glass NL (2016) Regulation of the lignocellulolytic response in filamentous fungi. *Fungal Biol Rev* 30:101-111.
7. Furukawa T*, et al.* (2009) Identification of specific binding sites for XYR1, a transcriptional activator of cellulolytic and xylanolytic genes in *Trichoderma reesei. Fungal Genet Biol* 46:564-574.
8. Hasper AA, Trindade LM, van der Veen D, van Ooyen AJ, & de Graaff LH (2004) Functional analysis of the transcriptional activator XlnR from *Aspergillus niger*. *Microbiol* 150:1367-1375.
9. Brunner K, Lichtenauer AM, Kratochwill K, Delic M, & Mach RL (2007) Xyr1 regulates xylanase but not cellulase formation in the head blight fungus *Fusarium graminearum*. *Curr Genet* 52:213-220.
10. Craig JP, Coradetti ST, Starr TL, & Glass NL (2015) Direct target network of the *Neurospora crassa* plant cell wall deconstruction regulators CLR-1, CLR-2, and XLR-1. *MBio* 6(5).
11. Thieme N*, et al.* (2017) The transcription factor PDR-1 is a multi-functional regulator and key component of pectin deconstruction and catabolism in *Neurospora crassa*. *Biotechnol Biofuels* 10:149.
12. Gruben BS*, et al.* (2014) *Aspergillus niger* RhaR, a regulator involved in L-rhamnose release and catabolism. *Appl Microbiol Biotechnol* 98:5531-5540.
13. Alazi E*, et al.* (2016) The transcriptional activator GaaR of *Aspergillus niger* is required for release and utilization of d-galacturonic acid from pectin. *FEBS Lett* 590:1804-1815.
14. Zhang L*, et al.* (2016) A novel Zn2 Cys6 transcription factor BcGaaR regulates D-galacturonic acid utilization in *Botrytis cinerea*. *Mol Microbiol* 100:247-262.

15. Battaglia E*, et al.* (2011) Analysis of regulation of pentose utilisation in *Aspergillus niger* reveals evolutionary adaptations in Eurotiales. *Stud Mycol* 69:31-38.
16. Klaubauf S, Zhou M, Lebrun MH, de Vries RP, & Battaglia E (2016) A novel L-arabinose-responsive regulator discovered in the rice-blast fungus *Pyricularia oryzae* (*Magnaporthe oryzae*). *FEBS Lett* 590:550-558.
17. Sun J & Glass NL (2011) Identification of the CRE-1 cellulolytic regulon in *Neurospora crassa*. *PLoS One* 6:e25654.
18. Ries LN, Beattie SR, Espeso EA, Cramer RA, & Goldman GH (2016) Diverse regulation of the CreA carbon catabolite repressor in *Aspergillus nidulans*. *Genetics* 203:335-352.
19. Xiong Y*, et al.* (2017) A fungal transcription factor essential for starch degradation affects integration of carbon and nitrogen metabolism. *PLoS Genet* 13:e1006737.
20. Nitta M*, et al.* (2012) A new Zn(II)(2)Cys(6)-type transcription factor BglR regulates beta-glucosidase expression in *Trichoderma reesei*. *Fungal Genet Biol* 49:388-397.
21. Xiong Y, Sun J, & Glass NL (2014) VIB1, a link between glucose signaling and carbon catabolite repression, is essential for plant cell wall degradation in *Neurospora crassa*. *PLoS Genet* 10:e1004500.
22. Ivanova C*, et al.* (2017) Genome sequencing and transcriptome analysis of *Trichoderma reesei* QM9978 strain reveals a distal chromosome translocation to be responsible for loss of *vib1* expression and loss of cellulase induction. *Biotechnol Biofuels* 10:209.
23. Znameroski EA*, et al.* (2012) Induction of lignocellulose-degrading enzymes in Neurospora crassa by cellodextrins. *Proc Natl Acad Sci USA* 109:6012-6017.
24. Gielkens MM, Dekkers E, Visser J, & de Graaff LH (1999) Two cellobiohydrolase-encoding genes from *Aspergillus niger* require D-xylose and the xylanolytic transcriptional activator XlnR for their expression. *Appl Environ Microbiol* 65:4340-4345.
25. Saier MH, Jr.*, et al.* (2016) The transporter classification database (TCDB): recent advances. *Nucleic Acids Res* 44:D372-379.
26. Langfelder P & Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559.
27. Coradetti ST*, et al.* (2012) Conserved and essential transcription factors for cellulase gene expression in ascomycete fungi. *Proc Natl Acad Sci USA* 109:7397-7402.
28. Benz JP*, et al.* (2014) A comparative systems analysis of polysaccharide-elicited responses in *Neurospora crassa* reveals carbon source-specific cellular adaptations. *Mol Microbiol* 91:275-299.
29. Benocci T*, et al.* (2018) ARA1 regulates not only l-arabinose but also d-galactose catabolism in *Trichoderma reesei*. *FEBS Lett* 592:60-70.
30. O'Malley RC*, et al.* (2016) Cistrome and epicistrome features shape the regulatory DNA landscape. *Cell* 165:1280-1292.
31. Derntl C*, et al.* (2013) Mutation of the Xylanase regulator 1 causes a glucose blind hydrolase expressing phenotype in industrially used *Trichoderma* strains. *Biotechnol Biofuels* 6:62.
32. Adnan M*, et al.* (2017) Carbon catabolite repression in filamentous fungi. *Int J Mol Sci* 19(1).

33. Espeso EA & Penalva MA (1994) In vitro binding of the two-finger repressor CreA to several consensus and non-consensus sites at the *ipnA* upstream region is context dependent. *FEBS Lett* 342:43-48.
34. Tamayo EN*, et al.* (2008) CreA mediates repression of the regulatory gene xlnR which controls the production of xylanolytic enzymes in *Aspergillus nidulans*. *Fungal Genet Biol* 45:984-993.
35. Wang B*, et al.* (2017) Identification and characterization of the glucose dual-affinity transport system in *Neurospora crassa*: pleiotropic roles in nutrient transport, signaling, and carbon catabolite repression. *Biotechnol Biofuels* 10:17.
36. Galazka JM*, et al.* (2010) Cellodextrin transport in yeast for improved biofuel production. *Science* 330:84-86.
37. Xiong Y*, et al.* (2014) The proteome and phosphoproteome of *Neurospora crassa* in response to cellulose, sucrose and carbon starvation. *Fungal Genet Biol* 72:21-33.
38. Li X*, et al.* (2015) Cellobionic acid utilization: from *Neurospora crassa* to *Saccharomyces cerevisiae*. *Biotechnol Biofuels* 8:120.
39. Li X*, et al.* (2015) Expanding xylose metabolism in yeast for plant cell wall conversion to biofuels. *eLife* 4.
40. Znameroski EA*, et al.* (2014) Evidence for transceptor function of cellodextrin transporters in *Neurospora crassa*. *J Biol Chem* 289:2610-2619.
41. Du J, Li S, & Zhao H (2010) Discovery and characterization of novel d-xylose-specific transporters from *Neurospora crassa* and *Pichia stipitis*. *Mol Biosyst* 6:2150-2156.
42. Li J, Lin L, Li H, Tian C, & Ma Y (2014) Transcriptional comparison of the filamentous fungus *Neurospora crassa* growing on three major monosaccharides D-glucose, D-xylose and L-arabinose. *Biotechnol Biofuels* 7:31.
43. Sloothaak J, Odoni DI, Martins Dos Santos VA, Schaap PJ, & Tamayo-Ramos JA (2016) Identification of a novel L-rhamnose uptake rransporter in the filamentous fungus *Aspergillus niger*. *PLoS Genet* 12:e1006468.
44. Dementhon K, Iyer G, & Glass NL (2006) VIB-1 is required for expression of genes necessary for programmed cell death in *Neurospora crassa*. *Eukaryot Cell* 5:2161-2173.
45. Xiang Q & Glass NL (2002) Identification of *vib-1*, a locus involved in vegetative incompatibility mediated by *het-c* in *Neurospora crassa. Genetics* 162:89-101.
46. Zhao J*, et al.* (2015) Identification of allorecognition loci in *Neurospora crassa* by genomics and evolutionary approaches. *Mol Biol Evol* 32:2417-2432.
47. Bok JW & Keller NP (2004) LaeA, a regulator of secondary metabolism in *Aspergillus* spp. *Eukaryot Cell* 3:527-535.
48. Huberman LB, Coradetti ST, & Glass NL (2017) Network of nutrient-sensing pathways and a conserved kinase cascade integrate osmolarity and carbon sensing in Neurospora crassa. *Proc Natl Acad Sci USA* 114:E8665-E8674.
49. Katz ME, Gray K-A, & Cheetham BF (2006) The *Aspergillus nidulans xprG* (*phoG*) gene encodes a putative transcriptional activator involved in the response to nutrient limitation. *Fungal Genet Biol* 43:190.
50. Mittenhuber G (2001) Phylogenetic analyses and comparative genomics of vitamin B6 (pyridoxine) and pyridoxal phosphate biosynthesis pathways. *J Mol Microbiol Biotechnol* 3:1-20.

51. Casado Lopez S*, et al.* (2018) Induction of genes encoding plant cell wall-degrading carbohydrate-active enzymes by lignocellulose-derived monosaccharides and cellobiose in the white-rot fungus D*ichomitus squalens*. *Appl Environ Microbiol* 84(11).
52. Zhang W*, et al.* (2013) Two major facilitator superfamily sugar transporters from *Trichoderma reesei* and their roles in induction of cellulase biosynthesis. *J Biol Chem*. 288:32861-72.
53. Brienzo M, Monte JR, & Milagres AM (2012) Induction of cellulase and hemicellulase activities of *Thermoascus aurantiacus* by xylan hydrolyzed products. *World J Microbiol Biotechnol* 28:113-119.
54. Schape P*, et al.* (2019) Updating genome annotation for the microbial cell factory *Aspergillus niger* using gene co-expression networks. *Nucleic Acids Res* 47:559-569.
55. Benz JP*, et al.* (2014) A comparative systems analysis of polysaccharide-elicited responses in *Neurospora crassa* reveals carbon source-specific cellular adaptations. *Mol Microbiol* 91:275-99.
56. Seiboth B & Metz B (2011) Fungal arabinan and L-arabinose metabolism. *Appl Microbiol Biotechnol* 89:1665-1673.
57. Mojzita D, Herold S, Metz B, Seiboth B, & Richard P (2012) L-xylo-3-hexulose reductase is the missing link in the oxidoreductive pathway for D-galactose catabolism in filamentous fungi. *J Biol Chem* 287:26010-26018.
58. Niu J*, et al.* (2015) The interaction of induction and repression mechanisms in the regulation of galacturonic acid-induced genes in *Aspergillus niger*. *Fungal Genet Biol* 82:32-42.
59. Haupt S, Louria-Hayon I, & Haupt Y (2003) P53 licensed to kill? Operating the assassin. *J Cell Biochem.* 88:76-82.
60. Xu L, Ajimura M, Padmore R, Klein C, & Kleckner N (1995) NDT80, a meiosis-specific gene required for exit from pachytene in *Saccharomyces cerevisiae*. *Mol Cell Biol* 15:6572-6581.
61. Hutchison EA & Glass NL (2010) Meiotic regulators Ndt80 and IME-2 have different roles in *Saccharomyces* and *Neurospora*. *Genetics* 185:1271-1282.
62. Katz ME (2019) Nutrient sensing-the key to fungal p53-like transcription factors? *Fungal Genet Biol* 124:8-16.
63. Min K, Biermann A, Hogan DA, & Konopka JB (2018) Genetic analysis of NDT80 family transcription factors in *Candida albicans* using new CRISPR-Cas9 approaches. *mSphere* 3(6).
64. Daskalov A, Heller J, Herzog S, Fleissner A, & Glass NL (2017) Molecular mechanisms regulating cell fusion and heterokaryon formation in filamentous fungi. *Microbiol Spectr* 5(2).
65. Roca MG, Weichert M, Siegmund U, Tudzynski P, & Fleissner A (2012) Germling fusion via conidial anastomosis tubes in the grey mould *Botrytis cinerea* requires NADPH oxidase activity. *Fungal Biol* 116:379-387.
66. Ishikawa FH, Souza EA, Read ND, & Roca MG (2010) Live-cell imaging of conidial fusion in the bean pathogen, *Colletotrichum lindemuthianum*. *Fungal Biol* 114:2-9.
67. Fischer-Harman V, Jackson KJ, Munoz A, Shoji JY, & Read ND (2012) Evidence for tryptophan being a signal molecule that inhibits conidial anastomosis tube fusion during colony initiation in *Neurospora crassa*. *Fungal Genet Biol* 49:896-902.

68. Bi Q, Wu D, Zhu X, & Gillian Turgeon B (2013) *Cochliobolus heterostrophus* Llm1 - a Lae1-like methyltransferase regulates T-toxin production, virulence, and development. *Fungal Genet Biol* 51:21-33.
69. Palmer JM*, et al.* (2013) Secondary metabolism and development is mediated by LlmF control of VeA subcellular localization in *Aspergillus nidulans*. *PLoS Genet* 9:e1003193.
70. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, & Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* 42:D490-495.
71. Kim D*, et al.* (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 14:R36.
72. Trapnell C*, et al.* (2013) Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat Biotechnol* 31:46-53.
73. Ruepp A*, et al.* (2004) The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. *Nucleic Acids Res* 32:5539-5545.
74. Zhang Y*, et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 9:R137.
75. Shannon P*, et al.* (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498-2504.
76. Krzywinski M*, et al.* (2009) Circos: an information aesthetic for comparative genomics. *Genome Res* 19:1639-1645.

PNAS
www.pnas.org

Supplementary Information for

The regulatory and transcriptional landscape associated with carbon utilization in a filamentous fungus

Vincent W. Wu[a,b], Nils Thieme[c,g], Lori B. Huberman[a,b], Axel Dietschmann[c,f], Elias Bleek[c], David J. Kowbel[a], Juna Lee[d], Sara Calhoun[d], Vasanth Singan[d], Anna Lipzen[d], Yi Xiong[a,b,h], Remo Monti[d], Matthew J. Blow[d], Ronan C. O'Malley[d], Igor V. Grigoriev[d,5], J. Philipp Benz[d], N. Louise Glass[a,b,e,1]

[a]Department of Plant and Microbial Biology, University of California, Berkeley, [b]Energy Biosciences Institute, University of California, Berkeley, [c]Holzforschung München, TUM School of Life Sciences Weihenstephan, Technical University of Munich, Freising, Germany, [d]Joint Genome Institute, Walnut Creek, CA [e]Environmental Genomics and Systems Biology, Lawrence Berkeley National Laboratory, Berkeley, CA

[1]Corresponding Author: N. Louise Glass
Lglass@berkeley.edu

This PDF file includes:

- Supplementary Materials and Methods
- Figures S1 to S5
- Tables S1, S2 and S5
- Legends for Table S3, Table S4 and Datasets S1 to S5
- SI References

Other supplementary materials for this manuscript include the following:

- Table S3 and Table S4
- Datasets S1 to S5

Materials and Methods

Strains, Growth Conditions and RNA extraction

N. crassa wild type (FGSC 2489) and gene deletion strains (Table S5) were obtained from the Fungal Genetics Stock Center (1)(www.fgsc.net). The double deletion strains *Δpdr-1*; *Δpdr-2* and *Δvib-1*; *Δclr-3* were created by crossing the respective single deletion strains, selection with hygromycin B (Invitrogen), and

confirmation by PCR. The ΔNCU02853 and ΔNCU08807 deletion strains were generated as in (2) For Fig. S6b, a ; Δ*cre-1* delete from (3) was used. The *ara-1* over-expression strain (OxNCU05414) was constructed by insertion of the *ara*-1 ORF into a vector having a glyceraldehyde-3-phosphate dehydrogenase (GPD) promoter and a cyclosporin resistance marker (4), which was introduced into a Δ*ara-1* strain.

For all RNAseq experiments, conidia obtained from 10-day-old pre-grown cultures were used to inoculate 3 ml of 1 x Vogel's salts (VMM) (5) plus 2% (w/v) sucrose at $1x10^6$ cells/mL in 24 well Whatman Uniplates. The bottoms of the wells for each Uniplate were initially scratched with a sharp needle to allow adherence and formation of mycelial mat. After 16 hrs of growth, mycelia were washed three times in 1 x VMM (5) without added carbon and subsequently transferred to 1 x VMM with a new carbon source. For induction conditions, 2mM mono and disaccharides were used, based on previously published reports for induction (6), while for complex polysaccharides and plant biomass 1% (w/v) was used. Plant biomass was obtained by grinding plant material to ~.08mm size. Carbon conditions and sources are included in Table S1**.** After 4 hrs of induction, mycelia were harvested over Whatman #1 filter paper and flash frozen in liquid nitrogen for storage at -80˚C. ~~Total RNA was isolated with TRIzol reagent (Invitrogen), treated with TURBO DNase (Thermo Fisher), and purified using Qiagen RNAeasy Purification Kit. RNA was tested for quality using agarose gel electrophoresis. Three biological replicates were obtained and analyzed for every condition and strain.~~

RNA extractions were performed on -80˚C stored biomass using TRIzol (Invitrogen) and chloroform. Entire biomass sample from each 3ml cultures were added to 1ml of TRIzol in a screwcap tube along with ~.5cm^2 of 0.1mm silica beads. Tubes containing biomass, TRIzol and beads were bead beaten for 30 seconds, then allowed to incubate at room temperature for 5 minutes on a rocker. 200ul of chloroform was added to each tube, vortexed, and centrifuged for phase separation. 400ul of aqueous phase from each sample was combined with 400ul isopropanol and incubated at room temperature for 10 minutes on rocker for RNA precipitation. Samples were centrifuged at 4˚C for 10 minutes. RNA pellets that formed were washed with 75% ethanol and centrifuged at 4˚C for 2 minutes. Ethanol was removed via pipet, and RNA pellet allowed to dry for several minutes with cap open. RNA pellet was resuspended in 40ul water, and treated with 2ul of Turbo DNAse (Thermo Fisher), in a 50ul reaction. After 20 minute incubation at 37˚C, RNA was cleaned up using Qiagen RNeasy Pufirication Kit. RNA was tested for quality using agarose gel electrophoresis. Three biological replicates were obtained and analyzed for every condition and strain.

RNA-seq

Stranded cDNA libraries were generated using the Illumina Truseq Stranded RNA LT kit. mRNA was purified from 1 µg of total RNA using magnetic beads containing poly-dT oligos. mRNA was fragmented and reverse transcribed using random hexamers and SSII (Invitrogen) followed by second strand synthesis. The fragmented cDNA was treated with end-pair, A-tailing, adapter ligation, and 10 cycles of PCR. The prepared library was then quantified using KAPA Biosystem's next-generation sequencing library qPCR kit and run on a Roche LightCycler 480 real-time PCR instrument. The quantified library was then multiplexed with other libraries, and the pool of libraries was then prepared for sequencing on the Illumina HiSeq sequencing platform utilizing a TruSeq paired-end cluster kit, v3, and Illumina's cBot instrument to generate a clustered flowcell for sequencing.

Sequencing of the flowcell was performed on the Illumina HiSeq2000 sequencer using a TruSeq SBS sequencing kit, v3, following a 2x100 indexed run recipe.

Filtered reads were mapped against *N. crassa* OR74A genome (v12) using Tophat 2.0.4 (7). Transcript abundance was estimated with Cufflinks 2.0.2 (8) in fragments per kilobase of transcript per million mapped reads (FPKMs) using upper quartile normalization and mapping against reference isoforms from the Broad Institute. Tophat mapped reads were additionally counted by HTSeq 0.6.0 (9) to obtain raw counts. Differential expression analysis was performed on raw counts with DEseq2 version 3.3 (10) using data from biological triplicates. Hierarchical clustering was performed using Python visualization library Seaborn (https://seaborn.pydata.org/). Parameters for clustering of PCWDE expression across 2mM sugar conditions: method = complete, metric = jaccard. Parameters for clustering DEseq fold change for VIB-1 direct targets: method = weighted. 'Hierarchical Clustering Explorer' v3.5 software (http:// www.cs.umd.edu/hcil/multi-cluster/hce3.html) was used to compare the mean FPKMs of the RNA-Seq libraries replicates (11, 12). A fold-change of $2^{1.5}$ (corresponding to 2.8 -fold change) was used because this cutoff used for these experiments included genes identified from previous RNA-seq experiments that were direct targets of CLR-2 and XLR-1 (13). Data is available here: https://genome.jgi.doe.gov/portal/TheFunENCproject/TheFunENCproject.info.html

Weighted Correlation Network Analysis (WGCNA) and FUNCAT analyses

After filtering genes out due to low expression (<10 FKPM) across >95% of all conditions, 6,742 genes were used in correlation analysis. The correlations were scaled using soft power of 9, assuming a scale-free network. Hierarchical clustering was applied to identify co-expressed gene modules with a minimum module size of 30 genes. The network with a force-directed layout was visualized using Cytoscape version 3.3 (Shannon et al., 2003). The scaled correlations between gene pairs above 0.15 are represented as edges in the network. Functional enrichment for the gene modules was performed using functional assignments in FunCatDB (14). The enrichment was calculated by the hypergeometric test and adjusted for multiple hypothesis testing using the Bonferroni correction. CAZy gene annotations were obtained from the JGI MycoCosm portal, Neucr2 (15).

MFS transporter tree construction

Putative transporters were predicted using data from the transporter classification database (TCDB) (http://www.tcb.org/) to generate a list of all *N. crassa* transporters and major facilitator superfamily (MFS) transporters were selected from this list. Protein sequences of all MFS genes were aligned using MAFFT version 7 (16) and used to construct a maximum likelihood phylogeny using RAxML (17). FigTree v1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/) was used for visualization.

Enzyme activity and transport assays

Strains were grown for 10 days in slants containing VMM (5) before conidia were harvested. 1 x 10^6 conidia/ml were transferred to 24 deep well plates containing 3 ml of 1x VMM plus the appropriate carbon source. Uptake assays were performed as described in (3, 18) with the following modifications. *N. crassa* WT, Δ*sut-28*, and Δ*cre-1* strains were pre-grown for 16 hr in 3 ml medium containing 1x VMM and 2% sucrose. The mycelial biomass was washed three times in 1x VMM solution and then transferred to an induction medium containing 1x VMM plus either

0.5% pectin, 0.5% pectin and 2% (w/v) D-glucose (D-glc), or 0.5% xylan. The samples were incubated for an additional 4 hrs and then washed again, as described above. WT and Δ*cre-1* strains were induced in 0.5% pectin or 0.5% pectin plus 2% D-glc and transferred to either 100 µM L-rha or 100 µM L-rha plus 100 µM D-glc as uptake solution. WT and Δ*sut-28* strains were transferred to uptake solutions containing either 100 µM L-rha, 90 µM D-fucose (D-fuc, VWR, A16789), 90 µM D-xyl or 90 µM D-galactose (D-gal, Sigma Aldrich, G0750).

Cellulase activity assays were modified from Coradetti *et al* (19). Briefly, 3 x 10^6 conidia of the indicated strains were inoculated into 3 ml VMM + 2% sucrose in round-bottomed, deep-well 24-well plates and grown at 25ºC in constant light with constant shaking at 200 rpm for 24 hr. The VMM + sucrose was vacuumed out of the well and the mycelial cell mass was washed in VMM lacking a carbon source, resuspended in media containing 2% Avicel as the sole carbon source, and incubated as described above. Culture supernatants were harvested 72 hrs post-transfer. Enzyme activity present in the culture supernatant was assayed with the Remazol brilliant Blue R-conjugated carboxymethyl cellulose kit (Megazyme).

Quantitative real-time PCR

Gene expression was analyzed by quantitative real-time PCR (RT-qPCR) as described in (12). The expression of *sut-28* was determined for the WT and Δ*cre-1* strains induced for 4 h on 1x VMM plus either 2 mM L-rha or 2 mM L-rhamnose together with 2% D-glc. Relative expression levels were calculated against the WT strain induced on 2 mM L-rhamnose. NCU09034 expression was analyzed for WT, Δ*pdr-1*, and Δ*sut-28* strains induced for 4 h on 1x VMM plus 2 mM L-rhamnose. The actin gene (NCU04173) was used to normalize expression data.

Genomic DNA library prep for DAP-seq

For genomic DNA isolation, the FGSC 2489 strain was grown on liquid VMM for 24 hours at 25˚C. Mycelia was filtered using Whatman #1 filter papers, and collected into 2ml tubes for flash freezing in liquid N_2 and cell rupturing. Cell rupturing was conducted by adding 1mm silica beads and along with DNA lysis buffer (0.05M NaOH, 1mM EDTA, 1% TritonX) and placed into bead beater for 1 minute. DNA was purified using DNeasy Blood & Tissue kit (Qiagen Inc.). DNA was sheared to 300 bp peak using Covaris LE220 sonicator. Size selection for sheared DNA was performed using AMPure XP beads to remove DNA above and below target molecular weight. Initially, sheared DNA was mixed in with AMPure XP beads (in PEG-8000) at a ratio of 100:60. At this ratio, beads bind DNA with molecular weight above 700 bp. Supernatant from this primary binding was taken and added to new beads where final ratio of DNA solution to PEG-8000 was at 100:90. At this ratio, DNA below ~300 bp does not bind to AMPure XP beads, and remaining DNA can be eluted for library preparation. The KAPA library kit for Illumina sequencing was used to prepare final libraries and stored at -20˚C for later use.

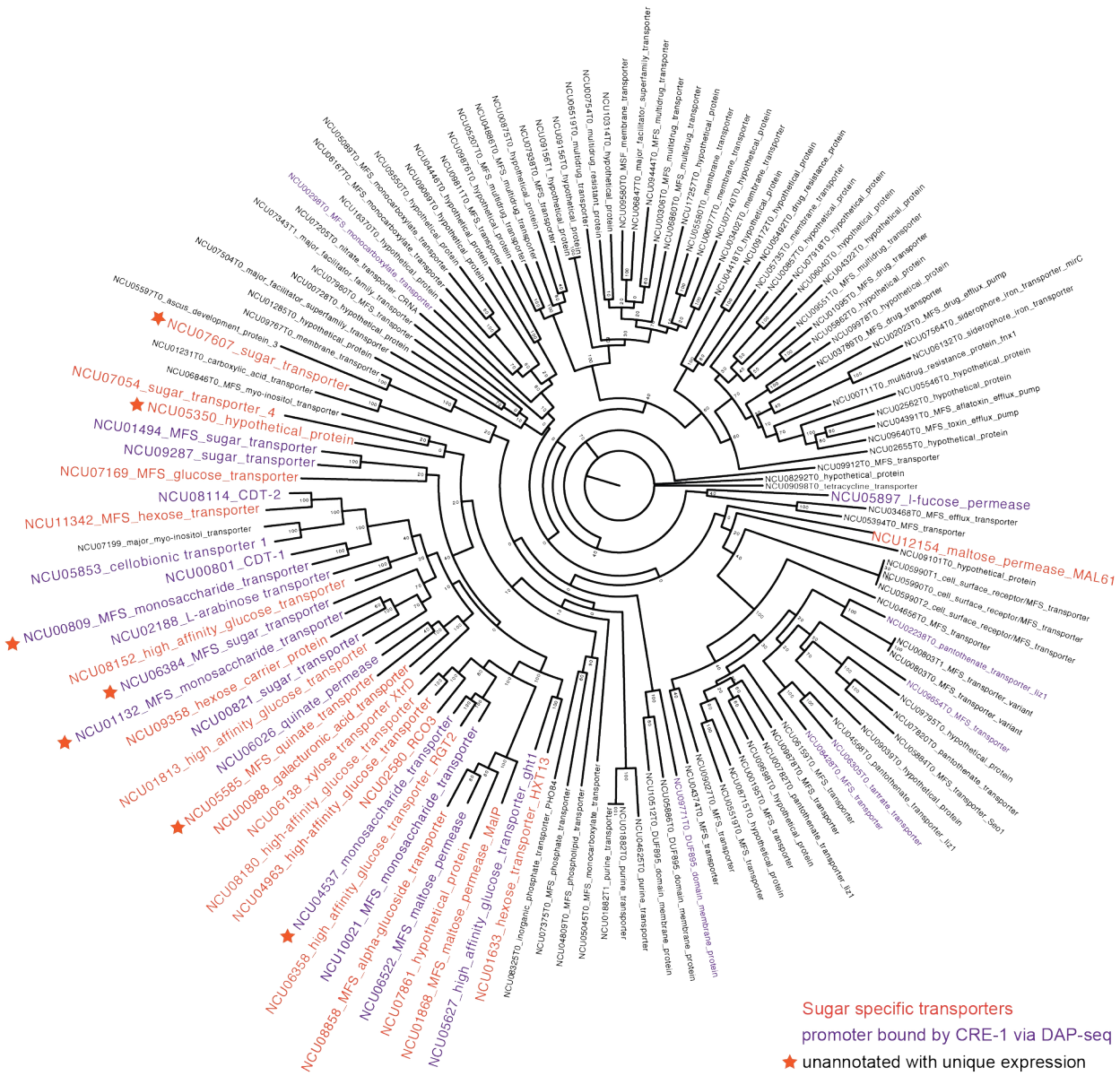
Transcription, translation, and DNA affinity purification (DAP)

Completed *in vitro* transcription and translation TnT reactions were incubated with 20ng of genomic DNA libraries, 1ug salmon sperm for blocking and 20ul Promega Magne HaloTag Beads on a rotator for 1 hr at room temperature. Bead bound proteins along with protein bound DNA were washed three times with 2.5% Tween20 in PBS. HaloTag beads were resuspended in 30ul ddH_2O and heated to 98˚C for 10 minutes to denature protein and release DNA fragments into solution.

Supernatant was transferred to a new tube for PCR amplification. DNA was amplified for final libraries using KAPA Hifi polymerase for 12-16 cycles of PCR to generate DAP-seq DNA libraries. A final DAP-seq DNA library was generated in the same conditions with no added plasmid into TnT Master Mix as a negative control. Single DAP-seq libraries were generated once for each transcription factor and sequenced with Illumina MiSeq 2x150BP runs.

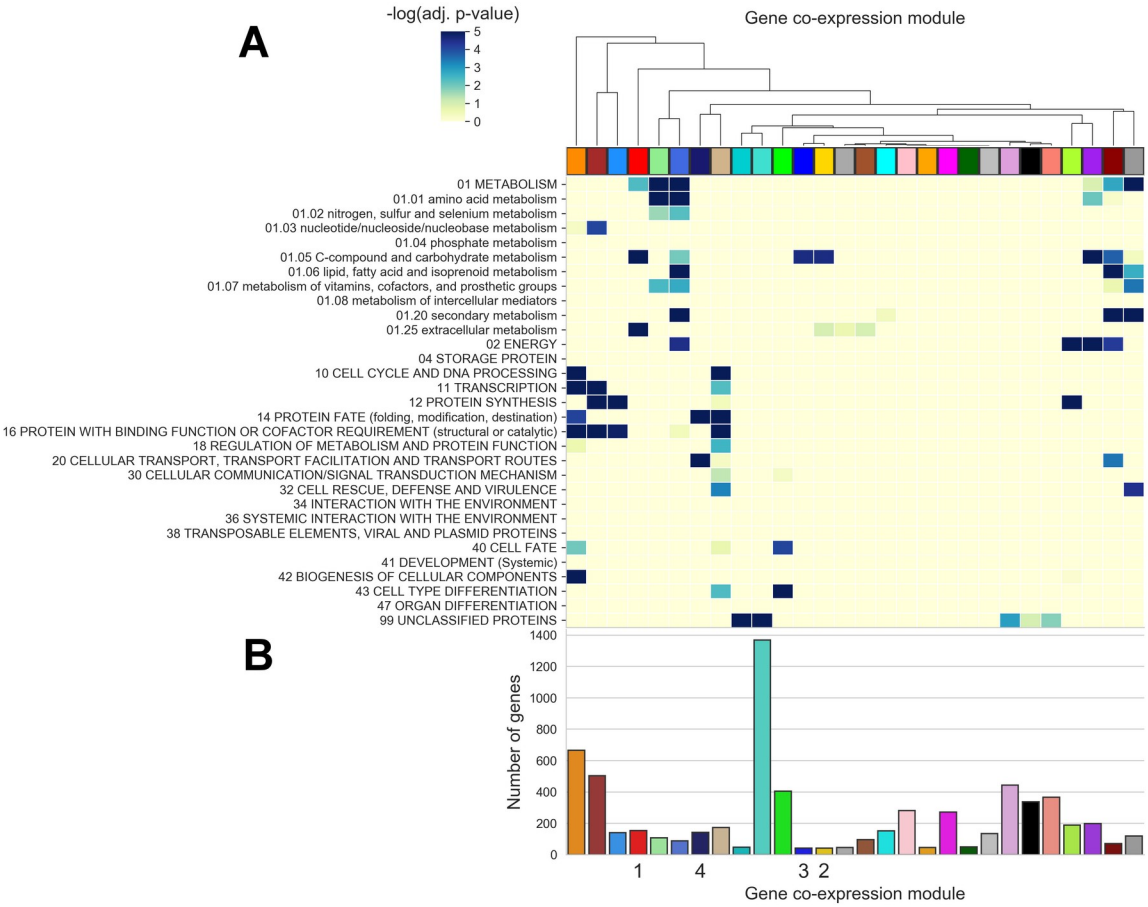
Motif construction

Motif discovery was performed using MEME v4.12.0 (20). Sequences of DAP-seq binding peaks were used as input for MEME motif discovery with flags maxw =20, minsites = 5, nmotifs = 8, denoting max width of motif, minimum number of sites for each motif and number of motifs to generate respectively. For CRE-1 motif discovery, all DAP-seq peaks within 3 kbp upstream of any translated genes were utilized as input sequences. For XLR-1, XLR^{A828V}, CLR-2 motif discovery, only DAP-seq peaks within 3 kbp upstream and from genes with $> 2^{1.5}$ fold reduction in Δtranscription factor RNA-seq data sets as compared to wild type were utilized for motif discovery. For VIB-1, DAP-seq peaks were within 1.5 kbp upstream for genes with $>2^{1.5}$ reduction in expression as compared to wild type for motif discovery.


NCU07607_sugar_transporter
NCU07054_sugar_transporter_4
NCU05350_hypothetical_protein
NCU01494_MFS_sugar_transporter
NCU09287_sugar_transporter
NCU07169_MFS_glucose_transporter
NCU08114_CDT-2
NCU11342_MFS_hexose_transporter
NCU05853_cellobionic transporter 1
NCU00801_CDT-1
NCU00809_MFS_monosaccharide_transporter
NCU02188_L-arabinose transporter
NCU08152_high_affinity_glucose_transporter
NCU06384_MFS_sugar_transporter
NCU09358_hexose_carrier_protein
NCU01132_MFS_monosaccharide_transporter
NCU00821_sugar_transporter
NCU06026_quinate_permease
NCU01813_high_affinity_glucose_transporter
NCU05585_MFS_quinate_transporter
NCU00988_galacturonic_acid_transporter
NCU06138_xylose transporter XtrD
NCU08180_high-affinity_glucose_transporter
NCU04963_high-affinity_glucose_transporter
NCU02580_RCO3
NCU04537_monosaccharide_transporter
NCU06358_high_affinity_glucose_transporter_RGT2
NCU10021_MFS_monosaccharide_transporter
NCU08858_MFS_alpha-glucoside_permease
NCU06522_MFS_maltose_transporter
NCU07861_hypothetical_protein
NCU01868_MFS_maltose_permease_MalP
NCU05627_high_affinity_glucose_transporter_ght1
NCU01633_hexose_transporter_HXT13
NCU05897_l-fucose_permease
NCU12154_maltose_permease_MAL61
Sugar specific transporters
promoter bound by CRE-1 via DAP-seq
unannotated with unique expression

1273

Fig. S1. MFS (Major Facilitator Superfamily) transporter maximum likelihood tree. Maximum likelihood tree built with RaxML (17). Bootstrap values denoted in the nodes of each branch. Tips are labeled with Gene ID and Broad v12 annotation (https://fungidb.org/fungidb/). Highlighted in red and purple are predicted or characterized sugar transporters. Marked with a red star are five transporters, which show increased expression across varied plant cell wall sugars at 2mM (SI Dataset 1). Highlighted in purple are transporter genes whose promoters are bound by CRE-1 via DAP-seq (SI Dataset 4).


A
-log(adj. p-value)
5
4
3
2
1
0
Gene co-expression module
01 METABOLISM
01.01 amino acid metabolism
01.02 nitrogen, sulfur and selenium metabolism
01.03 nucleotide/nucleoside/nucleobase metabolism
01.04 phosphate metabolism
01.05 C-compound and carbohydrate metabolism
01.06 lipid, fatty acid and isoprenoid metabolism
01.07 metabolism of vitamins, cofactors, and prosthetic groups
01.08 metabolism of intercellular mediators
01.20 secondary metabolism
01.25 extracellular metabolism
02 ENERGY
04 STORAGE PROTEIN
10 CELL CYCLE AND DNA PROCESSING
11 TRANSCRIPTION
12 PROTEIN SYNTHESIS
14 PROTEIN FATE (folding, modification, destination)
16 PROTEIN WITH BINDING FUNCTION OR COFACTOR REQUIREMENT (structural or catalytic)
18 REGULATION OF METABOLISM AND PROTEIN FUNCTION
20 CELLULAR TRANSPORT, TRANSPORT FACILITATION AND TRANSPORT ROUTES
30 CELLULAR COMMUNICATION/SIGNAL TRANSDUCTION MECHANISM
32 CELL RESCUE, DEFENSE AND VIRULENCE
34 INTERACTION WITH THE ENVIRONMENT
36 SYSTEMIC INTERACTION WITH THE ENVIRONMENT
38 TRANSPOSABLE ELEMENTS, VIRAL AND PLASMID PROTEINS
40 CELL FATE
41 DEVELOPMENT (Systemic)
42 BIOGENESIS OF CELLULAR COMPONENTS
43 CELL TYPE DIFFERENTIATION
47 ORGAN DIFFERENTIATION
99 UNCLASSIFIED PROTEINS
B
Number of genes
1400
1200
1000
800
600
400
200
0
1
4
3
2
Gene co-expression module

1284

1285

Fig. S2. Weighted gene co-expression network analysis heatmap and module information. (A) Heatmap of enrichment of general and metabolism-related functional categories by co-expressed gene module. Each column corresponds to a co-expressed gene module identified through weighted gene co-expression network analysis (WGCNA), with the same colors as the co-expressed gene modules shown in the network in Figure 1B. Each row corresponds to a functional category based on the classifications in FunCatDB (14). P-values reflect the significance of enrichment of genes assigned to a given functional category in a module and were calculated by the hypergeometric test. P-values were adjusted for multiple hypothesis testing using the Bonferroni correction. The color scale shows $-\log_{10}$ p-value with the most significant groups shown in dark blue and least significant in yellow. The order of the columns was determined by hierarchical clustering of modules on the $-\log_{10}$ p-values. (B) Number of genes in each co-expressed gene module from Figure 1B. See SI Dataset 2 for annotation of genes within the modules. ~~**Figure S2. Weighted gene co-expression network analysis heatmap and module information.** (A) Heatmap of enrichment of general and metabolism-related functional enrichment was performed using functional assignments in FunCatDB (14). Enrichment was calculated by the hypergeometric test and adjusted for multiple hypothesis testing using the Bonferroni correction. CAZyme gene annotations were obtained from the JGI MycoCosm portal, Neucr2 (15). The color scale shows $-\log_{10}$ p-value with the most significant groups shown in dark blue and least significant in yellow. (B) Number of genes in each co-expressed gene module from Figure 1B. See SI Dataset 2 for annotation of genes within the modules.~~

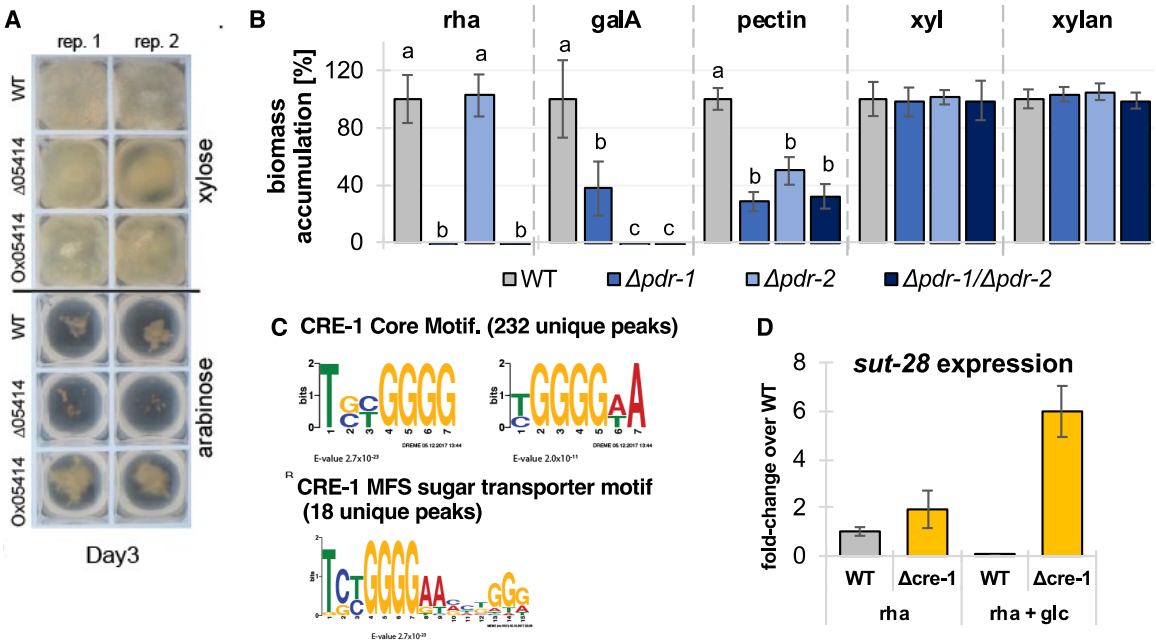


Fig.~~ure~~ S3. Growth of Δ*pdr* mutants and *ara-1* over-expression strain on different carbon sources, CRE-1 binding motifs and *sut-28* expression in a Δ*cre-1* mutant. (A) Biomass accumulation after shift to xylose or arabinose in WT, Δ*ara-1* (Δ05415) and *ara-1* over-expression (Ox0514) strains. (B) Relative biomass of FGSC2489 and Δ*sut-28* strain incubated in rhamnose, polygalacturonic acid, pectin, xylose, and xylan as determined by dry weight. (n=3) Significance was determined by ANOVA followed by a post-hoc Tukey's test. Different letters above bars represent significant difference ($p < 0.05$). (C) Top panel: Two CRE-1 binding motifs built using DREME v4.12.0 using all 232 DAP-seq peaks with the lowest E-value. Bottom panel: CRE-1 binding motifs built using MEME v4.12.0 using 18 DAP-seq binding peaks within the promoters of sugar transporters. (D) Relative expression of *sut-28* relative to *act* in FGSC2489 and Δ*cre-1* mutants. Strains were either induced in rhamnose or on rhamnose plus glucose (n >=3)
.

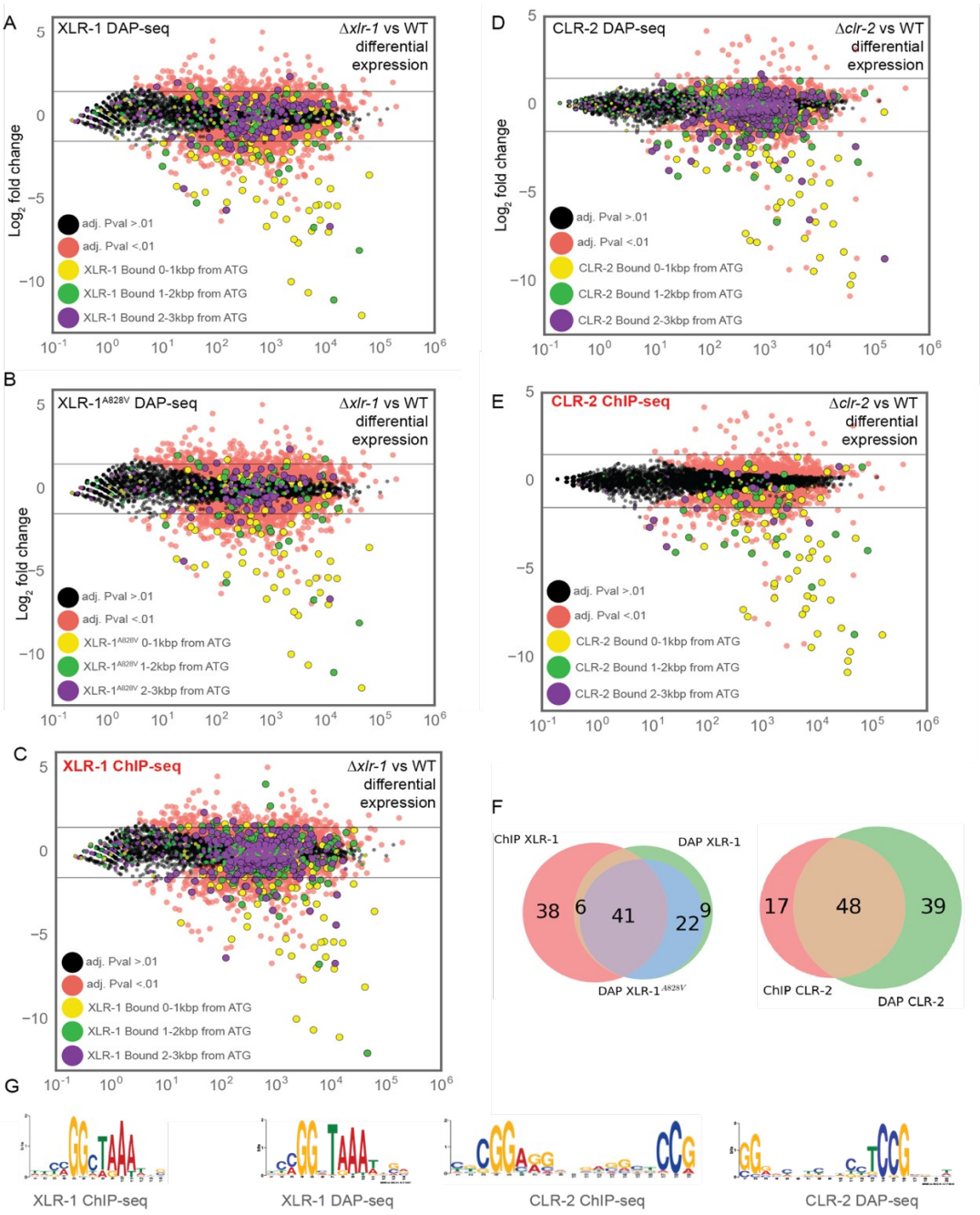
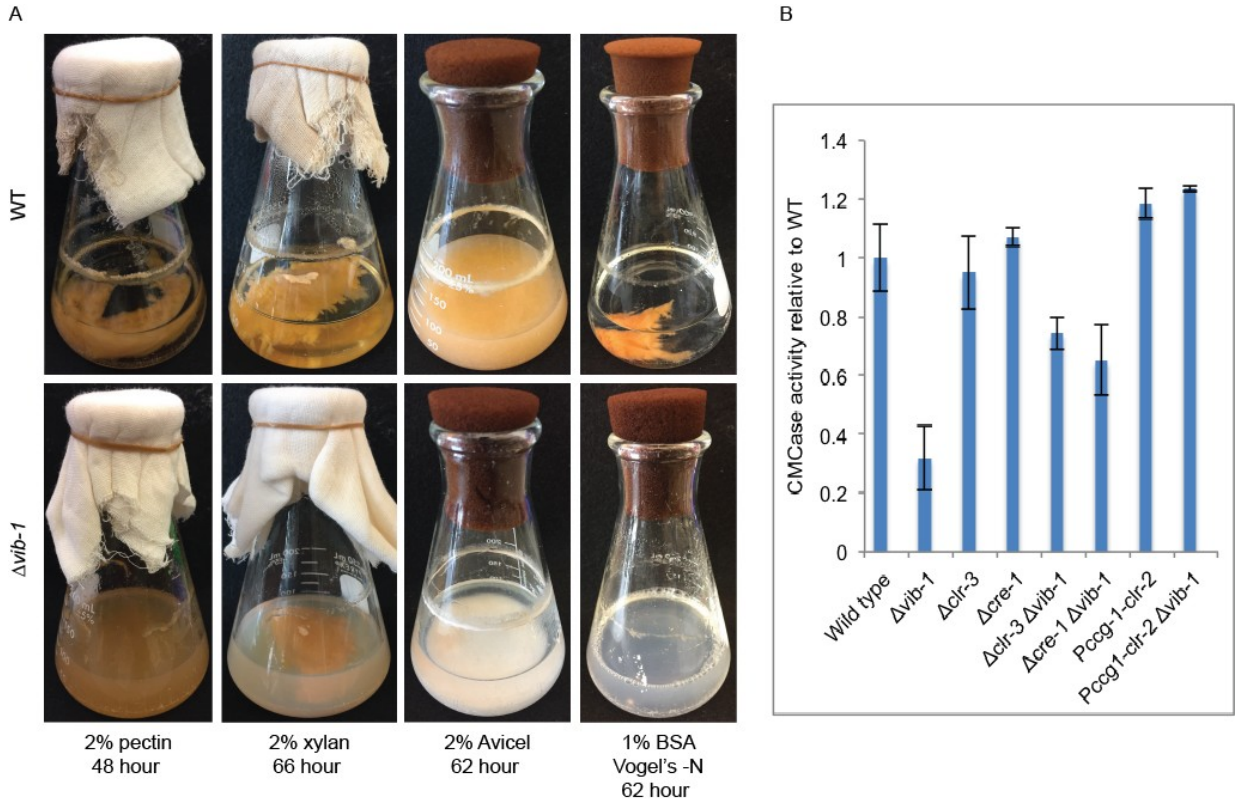


Fig. S4. DAP-seq validation utilizing published ChIP-seq data of XLR-1 and CLR-2. (A,-E) Genes with XLR-1 DAP-seq, XLR-1^{A828V} DAP-seq, XLR-1 ChIP-seq CLR-2 DAP-seq, and CLR2 ChIP-seq peaks overlaid onto differential expression data scatterplots of WT vs Δ*xlr-1* shifted to 1% xylan or WT vs Δ*clr-2* shifted to 1% Avicel, respectively. Gray lines represent chosen biological significance cutoff of $\pm 2^{1.5}$-fold change. Genes whose promoters are bound according to DAP-seq are highlighted in yellow, green or purple based on their distance from the ATG translation start site. "Base mean" is the mean of normalized counts for triplicates of both conditions tested. (F) Venn diagrams of direct targets from ChIP-seq (13) versus XLR-1 DAP-seq versus XLR-1^{A828V} DAP-seq and Venn diagram of CLR-2 direct targets as

discovered by ChIP-seq (13) versus DAP-seq. (G) XLR-1 and CLR-2 binding motifs built from binding peaks from differentially expressed genes whose promoters have binding peak sequences as discovered by ChIP-seq (13) versus DAP-seq. Motifs were built using DREME v4.12.0.



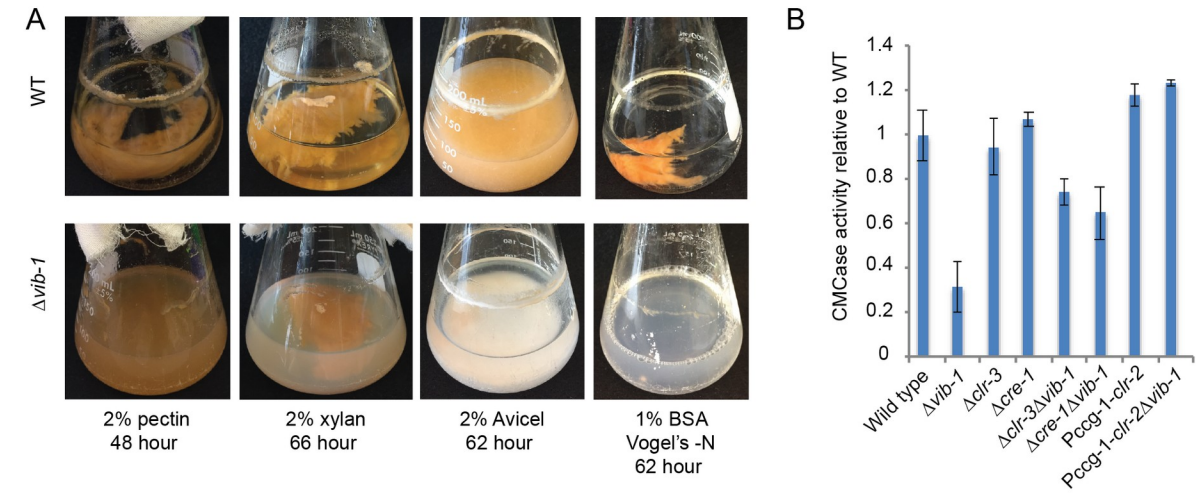


Fig. S5. Δ*vib-1* growth phenotypes on different carbon sources.
(A) Δ*vib-1* shows reduced growth and substrate depolymerization with pectin, xylan, or Avicel as the sole carbon source orBSA (as the sole carbon and nitrogen source). Cultures were directly inoculated from conidia at 25˚C at 200rpm, and pictures taken at the time at which WT cultures were able to depolymerize and clear the substrate. (B) CMCase activity of enzymes secreted into culture supernatants by the indicated strains relative to wild type 72h post-shift to media containing Avicel as the sole carbon source (p-adj<0.01, Student's *t*-test with Benjamini/Hochberg multiple hypothesis correction). P*ccg-1-clr-2* strains have constitutive expression of

clr-2, which results in inducer-independent induction of the CLR-2 cellulose regulon (21). (n >=3)

Table S1: Carbon sources used in this study

Condition	Source	CAS Number
sucrose	Sigma	57-50-1
fructose	Research organics	0609-06-03
xylose	Acros Organics 225990050	58-86-6
mannose	Acros D (+) 99+%	3458-38-4
maltose	Sigma D (+) min 98% <0.3% glucose <1.0% maltotriose	6363-53-7
arabinose	Sigma L (+) min 99% a3256	5328-37-0
cellobiose	Fluka D (+) >99%	
galactose	Acros D (+) 99+%	59-23-4
rhamnose	TCI R0013	10030-85-0
galacturonic acid	fluka 48280	
glucuronic acid	Sigma 98%	12/3/56
trehalose	Acros D 99%	6138-23-4
mannobiose	Megazyme	O-MBI
sorbose	Calbiochem L(-) 99.6%	CAS 87-79-6
glycerol	Fisher	56-81-5
ribose	TCI R0025	50-69-1
mannitol	Fisher	69-65-8
fucose	Sigma	2438-80-4
inulin	Sigma	9005-80-5
Avicel	Fluka	9004-34-6
xylan	Sigma from beechwood	9014-63-5
xyloglucan	Megazyme (from tamarind)	
galactomannan	Megazyme (carob low viscosity)	
glucomannan	Konjac Foods	
mixed linkage glucan	Megazyme (barley)	
pectin	Sigma from orange peel	9000-69-5
pectin esterified	Sigma	9046-40-6
polygalacturonic acid	Sigma	25990-10-7
rhamnogalacturonan	Megazyme (potato)	
arabinan	Megazyme (sugar beet)	
galactan	Megazyme (lupin)	
amylopectin	Sigma from corn A-7780	
amylose	Sigma from corn A-7043	
dioxane lignin	in house	
citrus peel	in house	
Miscanthus	in house	
corn stover	in house	
Wing Nut	in house	
Energy Cane	in house	
Switchgrass	in house	

Table S2. Predicted genes in the *Neurospora crassa* genome encoding plant biomass degrading enzymes and sugar transporters

Gene ID	Annotation	Gene name/enzyme family/TCBD[1]
NCU08412	beta-1,4-endomannanase	*gh5-7*
NCU00985	extracellular beta-mannosidase	*gh2-4*
NCU00890	intracellular beta-mannosidase	*gh2-1*
NCU08131	alpha-amylase	GH13
NCU09805	alpha-amylase	GH13
NCU05873	alpha-amylase	GH13
NCU05429	alpha-amylase	GH13
NCU06523	intracellular alpha-glucosidase	*gh13-4*
NCU07860	intracellular alpha-glucosidase	*gh13-5*
NCU03098	intracellular alpha-glucosidase	*gh15-1*
NCU00743	amylo-alpha-1,6-glucosidase	*gh13-7*
NCU01517	extracellular alpha-glucosidase *gla-1*	GH15
NCU02583	extracellular alpha-glucosidase	GH31
NCU09281	extracellular alpha-glucosidase	*gh31-1*
NCU04203	extracellular alpha-glucosidase	*gh31-2*
NCU04674	extracellular alpha-glucosidase	*gh31-3*
NCU04221	neutral trehalase	GH37
NCU08746	starch active LPMO	AA13
NCU00943	trehalase	GH37
NCU09170	alpha-arabinofuranosidase	*gh43-4*
NCU05965	alpha-arabinofuranosidase	*gh43-7*
NCU02343	alpha-arabinofuranosidase	*gh51-1*
NCU09775	alpha-arabinofuranosidase	*gh54-1*
NCU00642	extracellular beta-galactosidase	*gh35-1*
NCU04623	extracellular beta-galactosidase	*gh35-2*
NCU06861	glycosyl hydrolase family 43 protein	GH43
NCU01900	intracellular beta-xylosidase	*gh43-2*
NCU00972	endo-beta-1,4-galactanase	*gh53-1*
NCU05882	endo-beta-1,6-galactanase	*gh5-5*
NCU09702	endo-beta-1,6-galactanase	*gh5-6*
NCU00852	endoarabinanase	*gh43-1*
NCU02369	endogalacturonase	*gh28-1*
NCU06961	exogalacturonase	GH28
NCU09924	exoarabinanase	GH93
NCU00937	extracellular beta-glucuronidase	*gh79-1*
NCU09774	feruloyl esterase C	CE1
NCU09491	feruloyl esterase B *fae-1*	CE1
NCU08785	feruloyl esterase D faeD	CE1
NCU06326	pectate lyase *ply-1*	GH76
NCU08176	pectate lyase *ply-2*	PL3

NCU10045	pectin methyl esterase	*ce8-1*
NCU09976	rhamnogalacturonan acetyl esterase	*ce12-1*
NCU05598	rhamnogalacturonan lyase asd-1	PL4
NCU02654	unsaturated rhamnogalacturonyl hydrolase	*gh105-1*
NCU07351	alpha-glucuronidase	GH67
NCU06143	alpha-glucuronidase	GH67
NCU04885	alpha-xylosidase	GH31
NCU05924	beta-1,4-endoxylanase	*gh10-1*
NCU08189	beta-1,4-endoxylanase	*gh10-2*
NCU04997	beta-1,4-endoxylanase	*gh10-3*
NCU07130	beta-1,4-endoxylanase	*gh10-4*
NCU02855	beta-1,4-endoxylanase	*gh11-1*
NCU07225	beta-1,4-endoxylanase	*gh11-2*
NCU09923	extracellular beta-xylosidase	*gh3-7*
NCU00709	extracellular beta-xylosidase	*gh3-8*
NCU04870	acetyl xylan esterase	CE1
NCU04494	acetyl xylan esterase	CE1
NCU00710	acetyl xylan esterase	CE1
NCU05159	acetyl xylan esterase	CE1
NCU09663	acetyl xylan esterase	CE1
NCU09664	acetyl xylan esterase	CE1
NCU03181	acetyl xylan esterase	CE1
NCU00762	beta-1,4-endoglucanase	*gh5-1*
NCU05057	beta-1,4-endoglucanase	*gh7-1*
NCU04854	beta-1,4-endoglucanase	*gh7-2*
NCU04027	beta-1,4-endoglucanase	*gh7-3*
NCU05121	beta-1,4-endoglucanase	*gh45-1*
NCU05955	beta-1,4-endoglucanase	*gh74-1*
NCU00206	cellobiose dehydrogenase	CDH
NCU05923	cellobiose dehydrogenase	CDH
NCU03996	exo-beta-1,4-glucanase or cellobiohydrolase	*gh6-1*
NCU09680	exo-beta-1,4-glucanase or cellobiohydrolase	*gH6-2*
NCU07190	exo-beta-1,4-glucanase or cellobiohydrolase	*gh6-3*
NCU07340	cellobiohydrolase *cbh-1*	GH7
NCU05104	exo-beta-1,4-glucanase or cellobiohydrolase	*gh7-4*
NCU08760	polysaccharide monooxygenase 1 (PMO1)	AA9
NCU03328	polysaccharide monooxygenase 1 (PMO1)	AA9
NCU00836	polysaccharide monooxygenase 1 (PMO1)	AA9
NCU01867	polysaccharide monooxygenase 1 (PMO1)	AA9
NCU02344	polysaccharide monooxygenase 1 (PMO1)	AA9
NCU09764	polysaccharide monooxygenase 1 (PMO1)	AA9
NCU02240	polysaccharide monooxygenase 2 (PMO2)	AA9
NCU02916	polysaccharide monooxygenase 2 (PMO2)	AA9
NCU01050	polysaccharide monooxygenase 2 (PMO2)	AA9

NCU07760	polysaccharide monooxygenase 3 (PMO3)	AA9
NCU03000	polysaccharide monooxygenase 3 (PMO3)	AA9
NCU05969	polysaccharide monooxygenase 3 (PMO3)	AA9
NCU07520	polysaccharide monooxygenase 3 (PMO3)	AA9
NCU07898	polysaccharide monooxygenase 3 (PMO3)	AA9
NCU07974	polysaccharide monooxygenase 3 (PMO3)	AA9
NCU03641	extracellular beta-glucosidase	GH3
NCU08755	extracellular beta-glucosidase	GH3
NCU04952	extracellular beta-glucosidase	GH3
NCU00130	intracellular beta-glucosidase	GH1
NCU08054	intracellular beta-glucosidase	GH3
NCU05577	intracellular beta-glucosidase	GH3
NCU07487	intracellular beta-glucosidase	GH3
NCU09533	galacturonic acid reductase (GaaA)	
NCU07064	L-galactonate-dehdyratase (GaaB)	
NCU09532	L-threo-3-deoxy-hexulosonate adolase (GaaC)	
NCU01906	put. L-glyceraldehyde reductase (GaaD)	
NCU09035	L-rhamnose-1-dehydrogenase LRA1	
NCU03605	L-rhamnono-γ-lactonase LRA2	
NCU09034	rhamnonate dehydratase LRA3	
NCU05037	L- and L-KDR aldolase LRA4	
NCU00643	L-arabinitol-dehydrogenase *ard-1*	
NCU02188	L-arabinose transporter *lat-1*	
NCU00891	xylitol dehydrogenase	
NCU08384	xylose/arabinose reductase	
NCU09041	xylulose reductase	
NCU08687	galactokinase	
NCU04460	galactose-1-phosphate uridylyltransferase	
NCU08549	UDP-galactose 4-epimerase	
NCU07607	sugar transporter	2.A.1.12.2
NCU00450	sucrose transporter	2.A.2.6.1
NCU00801	cellodextrose transporter cdt-1	2.A.1.1.9
NCU00809	MFS monosaccharide transporter	2.A.1.1.9
NCU00821	sugar transporter	2.A.1.1.73
NCU00988	galacturonic acid transporter gat-1	2.A.1.1.7
NCU01132	MFS monosaccharide transporter	2.A.1.1.38
NCU01231	carboxylic acid transporter	2.A.1.12.2
NCU01494	MFS sugar transporter	2.A.1.1.69
NCU01633	hexose transporter HXT13	2.A.1.1.58
NCU01813	high affinity glucose transporter	2.A.1.1.39
NCU01868	MFS maltose permease MalP	2.A.1.1.10
NCU02188	L-arabinose transporter lat-1	2.A.1.1.39

NCU02582	sorbose-resistant-4	2.A.1.1.51
NCU04537	monosaccharide transporter	2.A.1.1.57
NCU04963	high-affinity glucose transporter	2.A.1.1.51
NCU05350	MFS transporter	2.A.1.1.40
NCU05585	MFS quinate transporter	2.A.1.1.7
NCU05627	high affinity glucose transporter ght1	2.A.1.1.36
NCU05853	cellobionic acid transporter cbt-1	2.A.1.1.9
NCU05897	l-fucose permease (rhamnose transporter)	2.A.1.7.1
NCU06026	quinate permease	2.A.1.1.7
NCU06138	xylose transporter XtrD aspergillus	2.A.1.1.40
NCU06358	high affinity glucose transporter RGT2	2.A.1.1.57
NCU06384	MFS sugar transporter	2.A.1.1.38
NCU06522	MFS maltose permease	2.A.1.1.10
NCU07054	sugar transporter 4	2.A.1.1.57
NCU07169	MFS glucose transporter	2.A.1.1.43
NCU07199	major myo-inositol transporter iolT	2.A.1.1.9
NCU07861	MFS transporter	2.A.1.1.10
NCU08114	cellodextrose transporter cdt-2	2.A.1.1.9
NCU08152	high affinity glucose transporter	2.A.1.1.39
NCU08180	high-affinity glucose transporter	2.A.1.1.68
NCU08858	MFS alpha-glucoside transporter	2.A.1.1.10
NCU09287	sugar transporter	2.A.1.1.69
NCU09321	sucrose transporter	2.A.2.6.1
NCU09358	hexose carrier protein	2.A.1.1.73
NCU10021	MFS monosaccharide transporter	2.A.1.1.57
NCU11342	MFS hexose transporter	2.A.1.1.9
NCU12154	maltose permease MAL61	2.A.1.1.11

[1]**TCDB:** The Transporter Classification Database (22).

Table S3. Predicted genes encoding DNA binding proteins in the *N. crassa* genome and carbon conditions used for testing expression of 34 transcription factors mutants

Table S4. Functional category (FunCAT) analyses of CRE-1-bound genes

Table S5. Strains used in this study

Name	Genotype	Source
Wild type	*Wild type mat A*	FGSC 2489 (1)
ΔNCU08042 (*clr-2*)	*Δclr-2::hygR mat A*	FGSC 15834
ΔNCU07705 (*clr-1*)	*Δclr-1::hygR mat A*	FGSC 11028
ΔNCU06971 (*xlr-1*)	*Δxlr-1::hygR mat A*	FGSC 11067
ΔNCU05414 (*ara-1*)	*ΔNCU05414 (ara-1)::hygR mat A*	FGSC 21219
ΔNCU04295 (*pdr-2*)	*ΔNCU04295 (pdr-2)::hygR mat a*	FGSC 18855
ΔNCU09033 (*pdr-1*)	*ΔNCU09033 (pdr-1)::hygR mat A*	FGSC 09033
ΔNCU03725 (*vib-1*)	*ΔNCU03725 (vib-1)::hygR mat A*	FGSC 11309
ΔNCU00282	*ΔNCU00282::hygR mat a*	FGSC 12566
ΔNCU00289	*ΔNCU00289::hygR mat A*	FGSC 11086
ΔNCU00808	*ΔNCU00808::hygR mat A*	FGSC 11123
ΔNCU10080	*ΔNCU10080::hygR mat A*	FGSC 17448
ΔNCU01074	*ΔNCU01074::hygR mat A*	FGSC 17482
ΔNCU01154 (*sub-1*)	*ΔNCU01154::hygR mat A*	FGSC 11126
ΔNCU01209	*ΔNCU01209::hygR mat a*	FGSC 17008
ΔNCU01312 (*rca-1*)	*ΔNCU01312::hygR mat A*	FGSC 11209
ΔNCU01386	*ΔNCU01386::hygR mat A*	FGSC 16372
ΔNCU01640 (*rpn-4*)	*ΔNCU01640::hygR mat A*	FGSC 11195
ΔNCU02203	*ΔNCU02203::hygR mat A*	FGSC 11884
ΔNCU02307	*ΔNCU02307::hygR mat a*	FGSC 11054
ΔNCU02853	*ΔNCU02853::hygR mat A*	this study
ΔNCU03421	*ΔNCU03421::hygR mat a*	FGSC 11149
ΔNCU03417	*ΔNCU03417::hygR mat A*	FGSC 21249
ΔNCU03643	*ΔNCU03643::hygR mat A*	FGSC 11049
ΔNCU03699	*ΔNCU03699::hygR mat a*	FGSC 11130
ΔNCU04058	*ΔNCU04058::hygR mat a*	FGSC 17238
ΔNCU04211	*ΔNCU04211::hygR mat a*	FGSC 11133
ΔNCU04848	*ΔNCU04848::hygR mat A*	FGSC 16718
ΔNCU04851	*ΔNCU04851::hygR mat a*	FGSC 11089
ΔNCU05024	*ΔNCU05024::hygR mat A*	FGSC 14730
ΔNCU05909	*ΔNCU05909::hygR mat a*	FGSC 11104
ΔNCU06173	*ΔNCU06173::hygR mat A*	FGSC 11366
ΔNCU06920	*ΔNCU06920::hygR mat A*	FGSC 17870
ΔNCU07728 (*sre-1*)	*ΔNCU07728::hygR mat a*	FGSC 11268
ΔNCU08055	*ΔNCU08055::hygR mat A*	FGSC 11269
ΔNCU08634	*ΔNCU08634::hygR mat a*	FGSC 20296
ΔNCU08899	*ΔNCU08899::hygR mat a*	FGSC 11048
ΔNCU09169	*ΔNCU09169::hygR mat A*	FGSC 19656
ΔNCU09252	*ΔNCU09252::hygR mat A*	FGSC 11394
ΔNCU10697	*ΔNCU10697::hygR mat A*	FGSC 21637
ΔNCU05897	*ΔNCU05897::hygR mat A*	FGSC 13717
Δpdr-1Δpdr-2	*ΔNCU09033::hygR; ΔNCU04295::hygR mat a*	this study
Δcre-1	*Δcre-1::hygR mat A*	this study
OxNCU05414 (*ara-1*)	*pgpd-1-NCU05414:csr-1; ΔNCU05414::hygR mat A*	this study
SC4B5	*Δclr-3::hyg^R mat A*	FGSC 14350 (2)
SC1B3	*Δvib-1::hyg^R mat a*	FGSC 11308 (2)
LHN694	*P_{ccg-1}-clr-2::his-3 Δsad-1::hyg^R rid-*	(21)

	1^{-}; Δ*clr-2::hyg*^{R} *mat A*	
LHN971	Δ*vib-1::hyg*^{R}; Δ*clr-3::hyg*^{R} *mat a*	This study
LHN972	Δ*vib-1::hyg*^{R}; Δ*clr-3::hyg*^{R} *mat A*	This study
LHN898	Δ*cre-1::hyg*^{R} *mat A*	(3)
VW205	Δ*cre-1::hyg*^{R}; Δ*vib-1::hyg*^{R} *mat a*	(23)
VW206	P_{ccg-1}-*clr-2*::*his-3* Δ*sad-1::hyg*^{R} *rid-1*$^{-}$; Δ*clr-2::hyg*^{R}; Δ*vib-1::hyg*^{R} *mat A*	(23)

1378
1379

SI Dataset 1. Normalized FPKM counts of wild type cells and selected mutants exposed to carbon conditions (Table S1)

SI Dataset 2. Lists of genes within 28 modules from WCGNA analysis

SI Dataset 3. DE-seq analysis of expression profiles of Δ*clr-1, Δclr-2, Δxlr-1,* Δ*pdr-2* (NCU04295), Δ*ara-1* (NCU05414) and Δ*vib-1* transcription factor mutants on various carbon sources

SI Dataset 4. DAP-seq data on XLR-1, XLR-1^{A828V}, CLR-2, CRE-1, VIB-1, and reanalyzed data for ChIP-seq data for XLR-1 and CLR-2

SI Dataset 5. The VIB-1 direct targets and core VIB-1 regulon based on RNA-seq and DAP-seq data

References NOT UPDATED

1. McCluskey K (2003) The Fungal Genetics Stock Center: from molds to molecules. *Adv. Appl. Microbiol.* 52:245-262.
2. Colot HV*, et al.* (2006) A high-throughput gene knockout procedure for *Neurospora* reveals functions for multiple transcription factors. *Proc Natl Acad Sci U S A* 103:10352-10357.
3. Huberman LB, Coradetti ST, & Glass NL (2017) Network of nutrient-sensing pathways and a conserved kinase cascade integrate osmolarity and carbon sensing in *Neurospora crassa*. *Proc Natl Acad Sci U S A* 114:E8665-E8674.
4. Bardiya N & Shiu PK (2007) Cyclosporin A-resistance based gene placement system for *Neurospora crassa*. *Fungal Genet Biol* 44:307-314.
5. Vogel HJ (1956) A convenient growth medium for *Neurospora* (medium N). *Microbial Genetics Bulletin* 13:42-43.
6. Znameroski EA*, et al.* (2012) Induction of lignocellulose-degrading enzymes in *Neurospora crassa* by cellodextrins. *Proc Natl Acad Sci USA* 109:6012-6017.
7. Kim D*, et al.* (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 14:R36.
8. Trapnell C*, et al.* (2013) Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat Biotechnol* 31:46-53.
9. Anders S, Pyl PT, & Huber W (2015) HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31:166-169.
10. Love MI, Huber W, & Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550.
11. Thieme N*, et al.* (2017) The transcription factor PDR-1 is a multi-functional regulator and key component of pectin deconstruction and catabolism in *Neurospora crassa*. *Biotechnol Biofuels* 10:149.
12. Benz JP*, et al.* (2014) A comparative systems analysis of polysaccharide-elicited responses in *Neurospora crassa* reveals carbon source-specific cellular adaptations. *Mol Microbiol* 91:275-299.
13. Craig JP, Coradetti ST, Starr TL, & Glass NL (2015) Direct target network of the *Neurospora crassa* plant cell wall deconstruction regulators CLR-1, CLR-2, and XLR-1. *MBio* 6(5).
14. Ruepp A*, et al.* (2004) The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. *Nucleic Acids Res* 32:5539-5545.
15. Grigoriev IV*, et al.* (2014) MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Res* 42:D699-704.
16. Katoh K, Misawa K, Kuma K, & Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 30:3059-3066.
17. Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688-2690.
18. Ha SJ*, et al.* (2011) Engineered *Saccharomyces cerevisiae* capable of simultaneous cellobiose and xylose fermentation. *Proc Natl Acad Sci USA* 108:504-509.
19. Coradetti ST*, et al.* (2012) Conserved and essential transcription factors for cellulase gene expression in ascomycete fungi. *Proc Natl Acad Sci USA* 109:7397-7402.

20. Bailey TL*, et al.* (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res* 37:W202-208.
21. Coradetti ST, Xiong Y, & Glass NL (2013) Analysis of a conserved cellulase transcriptional regulator reveals inducer-independent production of cellulolytic enzymes in *Neurospora crassa*. *Microbiologyopen* 2:595-609.
22. Saier MH, Jr.*, et al.* (2016) The Transporter Classification Database (TCDB): recent advances. *Nucleic Acids Res* 44:D372-379.
23. Xiong Y, Sun J, & Glass NL (2014) VIB1, a link between glucose signaling and carbon catabolite repression, is essential for plant cell wall degradation in *Neurospora crassa*. *PLoS Genet* 10:e1004500.